Contents

Original Papers

Digital Peer-Supported App Intervention to Promote Physical Activity Among Community-Dwelling Older Adults: Nonrandomized Controlled Trial (e56184)
Kento Tabira, Yuko Oguma, Shota Yoshihara, Megumi Shibuya, Manabu Nakamura, Natsue Doihara, Akihiro Hirata, Tomoki Manabe. .............................. 19

Designing Telemedicine for Older Adults With Multimorbidity: Content Analysis Study (e52031)
Nida Buawangpong, Kanokporn Pinyopornpanish, Suphawita Pliannuom, Nopakoon Nantsupawat, Nutchar Wiwat kunupakarn, Chaisiri Angkurawanon, Wichuda Jiraporncharoen. .................................................. 34

Examining the Effect of Contactless Intergenerational Befriending Intervention on Social Isolation Among Older Adults and Students’ Attitude Toward Companionship: Content Analysis (e47908)
Keya Sen, Nida Laheji, Zo Ramamonjarivelo, Cecil Renick, Randall Osborne, Brad Beauvais. ................................................................. 47

Determinants of Implementing an Information and Communication Technology Tool for Social Interaction Among Older People: Qualitative Content Analysis of Social Services Personnel Perspectives (e43999)
Johanna Fritz, Petra von Heideken Wåger, Annelie Gusdal, Rose-Marie Johansson-Pajala, Caroline Eklund. ............................................................. 58

Factors That Influence Successful Adoption of Real-Time Location Systems for Use in a Dementia Care Setting: Mixed Methods Study (e45978)
Lynn Haslam-Larmer, Alisa Grigorovich, Leia Shum, Andria Bianchi, Kristine Newman, Andrea Iaboni, Josephine McMurray. .............................. 72

Using the TrueLoo Smart Device to Record Toileting Sessions in Older Adults: Retrospective Validation and Acceptance Study (e50856)
Jordan Glenn, Parmoon Sarmadi, Paul Cristman, Gabrielle Kim, Ting-Hsuan Lin, Vikram Kashyap. ................................................................. 86

Efficacy of COMPAs, an App Designed to Support Communication Between Persons Living With Dementia in Long-Term Care Settings and Their Caregivers: Mixed Methods Implementation Study (e47565)
Ana Ansaldo, Michèle Masson-Trottier, Barbara Delacourt, Jade Dubuc, Catherine Dubé. ................................................................. 97

Digital Storytelling Intervention for Enhancing the Social Participation of People With Mild Cognitive Impairment: Co-Design and Usability Study (e54138)
Di Zhu, Abdullah Al Mahmud, Wei Liu. ................................................................. 155

Experiences of Older Adults, Physiotherapists, and Aged Care Staff in the TOP UP Telephysiotherapy Program: Interview Study of the TOP UP Interventions (e53010)
Rik Dawson, Heidi Gilchrist, Marina Pinheiro, Karn Nelson, Nina Bowes, Cathie Sherrington, Abby Haynes. ................................................................. 175
Acceptance of a Digital Assistant (Anne4Care) for Older Adult Immigrants Living With Dementia: Qualitative Descriptive Study (e50219)
Marloes Bults, Catharina van Leersum, Theodorus Olthuis, Egbert Siebrand, Zohrah Malik, Lili Liu, Antonio Miguel-Cruz, Jan Jukema, Marjolein den Ouden. .............................................................................. 191

Technology Usability for People Living With Dementia: Concept Analysis (e51987)
Shao-Yun Chien, Oleg Zaslavsky, Clara Berridge. .............................................................................................................. 202

The Prevalence of Missing Incidents and Their Antecedents Among Older Adult MedicAlert Subscribers: Retrospective Descriptive Study (e58205)
Antonio Miguel-Cruz, Hector Perez, Yoojin Choi, Emily Rutledge, Christine Daum, Lili Liu. ................................. 239

Refining Cultural Adaptations of a Behavioral Intervention for Latino Caregivers of People Living With Dementia: Qualitative Interview Study in Washington State (e53671)
Celeste Garcia, Miriana Duran, Magaly Ramirez. .............................................................................................................. 255

Expectation, Attitude, and Barriers to Receiving Telehomecare Among Caregivers of Homebound or Bedridden Older Adults: Qualitative Study (e48132)
Pansiree Onseng, Wichuda Jiraporncharoen, Sasiwimon Moonkayaow, Pimchai Veerasirikul, Nutchar Wiwatkinupakarn, Chaisiri Angkurawaranon, Kanokporn Pinypornpanich. ........................................................................ 268

A Web-Based Intervention Based on Acceptance and Commitment Therapy for Family Caregivers of People With Dementia: Mixed Methods Feasibility Study (e53489)
Golnaz. Atefi, Rosalia van Knippenberg, Sara Bartels, Andrés Losada-Baltar, María Márquez-González, Frans Verhey, Marjolein de Vugt. ................................. 279

The Oleracare Digital Caregiving Assistance Platform for Dementia Caregivers: Preliminary Evaluation Study (e55132)
Qiping Fan, Minh-Nguyet Hoang, Logan DuBose, Marcia Ory, Jeswin Vennatt, Diana Salha, Shinduk Lee, Tokunbo Falahun. ................................................. 298

Digital Adoption by an Organization Supporting Informal Caregivers During COVID-19 Pandemic Showing Impact on Service Use, Organizational Performance, and Carers’ Well-Being: Retrospective Population-Based Database Study With Embedded User Survey (e46414)
Ala Szczepura, Amir Khan, Deidre Wild, Sara Nelson, Sonja Woodhouse, Mark Collinson. .................................................. 332

Sleep Duration and Functional Disability Among Chinese Older Adults: Cross-Sectional Study (e53548)
Minjing Luo, Yue Dong, Bingbing Fan, Xinyue Zhang, Hao Liu, Changhao Liang, Hongguo Rong, Yutong Fei. ................................................................. 348

Adoption and Use of Telemedicine and Digital Health Services Among Older Adults in Light of the COVID-19 Pandemic: Repeated Cross-Sectional Analysis (e52317)
Motti Haimi, Ruslan Sergienko. ................................................................................................................................................. 407

Experiences of a Community-Based Digital Intervention Among Older People Living in a Low-Income Neighborhood: Qualitative Study (e52292)
Si Lu, Sungwon Yoon, Wan Yee, Nerice Heng Wen Ngiam, Kennedy Ng, Lian Low. ................................................................. 428

Digital Literacy Training for Low-Income Older Adults Through Undergraduate Community-Engaged Learning: Single-Group Pretest-Posttest Study (e51675)
Lisa Miller, Rachel Callegari, Theresa Abah, Helen Fann. ........................................................................................................... 441

Long-Term Adoption of Televits in Nursing Homes During the COVID-19 Crisis and Following Up Into the Postpandemic Setting: Mixed Methods Study (e55471)
Tobias Martin, Sarah Veldeman, Heidrun Großmann, Paul Fuchs-Frohnhofen, Michael Czaplik, Andreas Follmann. .......................................................... 456
Automatic Spontaneous Speech Analysis for the Detection of Cognitive Functional Decline in Older Adults: Multilanguage Cross-Sectional Study (e50537)
Emilia Ambrosini, Chiara Giangregorio, Eugenio Lomurzo, Sara Moccia, Marios Milis, Christos Loizou, Domenico Azzolino, Matteo Cesari, Manuel Cid Gala, Carmen Galán de Isla, Jonathan Gomez-Raja, Nunzio Borghese, Matteo Matteucci, Simona Ferrante. 837

Promoting Personalized Reminiscence Among Cognitively Intact Older Adults Through an AI-Driven Interactive Multimodal Photo Album: Development and Usability Study (e49415)
Xin Wang, Juan Li, Tianyi Liang, Wordh Hasan, Kimia Zaman, Yang Du, Bo Xue, Cui Tao. 874

Reviews

Characteristics and Range of Reviews About Technologies for Aging in Place: Scoping Review of Reviews (e50286)
Jenny Bergschöld, Mari Gunnes, Arne Eide, Eva Lassemo. 218

The Use of Digital Technologies in the Promotion of Health Literacy and Empowerment of Informal Caregivers: Scoping Review (e54913)
Suzete Soares, Louise Hoffmeister, Maria Fernandes, Adriana Henriques, Andreia Costa. 317

The Implementation Success of Technology-Based Counseling in Dementia Care: Scoping Review (e51544)
Dorothee Bauernschmidt, Janina Wittmann, Julian Hirt, Gabriele Meyer, Anja Bieber. 753

Patient and Public Involvement in Technology-Related Dementia Research: Scoping Review (e48292)
Pippa Kirby, Helen Lai, Sophie Horrocks, Matthew Harrison, Danielle Wilson, Sarah Daniels, Rafael Calvo, David Sharp, Caroline Alexander. 7 9

Strategies to Mitigate Age-Related Bias in Machine Learning: Scoping Review (e53564)
Charlene Chu, Simon Donato-woodger, Shehroz Khan, Tianyu Shi, Kathleen Leslie, Samira Abbasgholizadeh-Rahimi, Rune Nyrop, Amanda Grenier. 782

Corrigenda and Addenda

Correction: Detecting Anomalies in Daily Activity Routines of Older Persons in Single Resident Smart Homes: Proof-of-Concept Study (e58394)
Zahraa Shahid, Saguna Saguna, Christer Åhlund. 815
Characterizing Walking Behaviors in Aged Residential Care Using Accelerometry, With Comparison Across Care Levels, Cognitive Status, and Physical Function: Cross-Sectional Study

Riona Mc Ardle¹,²,*, PhD; Lynne Taylor³,*, PhD; Alana Cavadinò³, PhD; Lynn Rochester¹,²,4, PhD; Silvia Del Din¹,², PhD; Ngaire Kerse³, PhD

Abstract

Background: Walking is important for maintaining physical and mental well-being in aged residential care (ARC). Walking behaviors are not well characterized in ARC due to inconsistencies in assessment methods and metrics as well as limited research regarding the impact of care environment, cognition, or physical function on these behaviors. It is recommended that walking behaviors in ARC are assessed using validated digital methods that can capture low volumes of walking activity.

Objective: This study aims to characterize and compare accelerometry-derived walking behaviors in ARC residents across different care levels, cognitive abilities, and physical capacities.

Methods: A total of 306 ARC residents were recruited from the Staying UpRight randomized controlled trial from 3 care levels: rest home (n=164), hospital (n=117), and dementia care (n=25). Participants’ cognitive status was classified as mild (n=87), moderate (n=128), or severe impairment (n=61); physical function was classified as high-moderate (n=74) and low-very low (n=222) using the Montreal Cognitive Assessment and the Short Physical Performance Battery cutoff scores, respectively. To assess walking, participants wore an accelerometer (Axivity AX3; dimensions: 23×32.5×7.6 mm; weight: 11 g; sampling rate: 100 Hz; range: ±8 g; and memory: 512 MB) on their lower back for 7 days. Outcomes included volume (ie, daily time spent walking, steps, and bouts), pattern (ie, mean walking bout duration and alpha), and variability (of bout length) of walking. Analysis of covariance was used to assess differences in walking behaviors between groups categorized by level of care, cognition, or physical function while controlling for age and sex. Tukey honest significant difference tests for multiple comparisons were used to determine where significant differences occurred. The effect sizes of group differences were calculated using Hedges' g (0.2-0.4: small, 0.5-0.7: medium, and 0.8: large).

Results: Dementia care residents showed greater volumes of walking (P<.001; Hedges g=1.0-2.0), with longer (P<.001; Hedges g=0.7-0.8), more variable (P=.008 vs hospital; P<.001 vs rest home; Hedges g=0.6-0.9) bouts compared to other care levels with a lower alpha score (vs hospital: P>.001; Hedges g=0.9, vs rest home: P=.004; Hedges g=0.8). Residents with severe cognitive impairment took longer (P<.001; Hedges g=0.5-0.6), more variable (P<.001; Hedges g=0.4-0.6) bouts, compared to those with mild and moderate cognitive impairment. Residents with low-very low physical function had lower walking volumes (total walk time and bouts per day: P<.001; steps per day: P=.005; Hedges g=0.4-0.5) and higher variability (P=.04; Hedges g=0.2) compared to those with high-moderate capacity.

Conclusions: ARC residents across different levels of care, cognition, and physical function demonstrate different walking behaviors. However, ARC residents often present with varying levels of both cognitive and physical abilities, reflecting their complex multimorbid nature, which should be considered in further work. This work has demonstrated the importance of considering a nuanced framework of digital outcomes relating to volume, pattern, and variability of walking behaviors among ARC residents.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12618001827224; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=376298&isReview=true

(JMIR Aging 2024;7:e53020) doi:10.2196/53020
KEYWORDS

residential aged care facility; cognitive dysfunction; mobility limitation; accelerometry; physical activity; aged residential care

Introduction

Physical mobility, such as walking, is a key predictor of health [1] and is considered a multifaceted experience that interconnects the physical, mental, social, and emotional needs of an individual with their sense of self [2,3]. Loss of physical mobility (eg, reduced volume of walking) is associated with increased safety risks (eg, falls), social withdrawal, and poorer well-being [4,5]. Supporting residents’ physical mobility in aged residential care (ARC) can decelerate the progression of disabilities and dependency [6]. ARC refers to long-term full-time residential care, which provides multiple levels of care depending on an individual’s needs. Other common terms for ARC include assisted living facilities, care homes, and nursing homes. It is recommended that all residents who can ambulate, regardless of cognitive abilities, should increase their activity levels to support their functional independence [7]. Continuous remote digital monitoring of mobility outcomes has been proposed as a method to objectively quantify changes in walking behaviors. This approach will inform the development of interventions aimed at better supporting mobility, which is a key factor in influencing well-being and function in the older population [2].

Accelerometers are the most common method to continuously assess walking behaviors in ARC residents [8], with outcomes relating to volume (eg, steps per day) and intensity (eg, moderate-vigorous physical activity) of walking activities most frequently reported. Based on the current literature, ARC residents primarily participate in low volumes of light-intensity walking and show little variation in their walking behaviors [8]. Based on prevailing gaps in the literature, current recommendations for the assessment of walking behaviors in ARC include the use of validated digital methods that can capture very low volumes of activity, using low cutoff thresholds (eg, any walking activity ≥3 steps), and derive standardized outcomes relating to volume (ie, the amount or duration of walking activity), pattern (ie, the distribution of walking activity across a time period), and variability (ie, changes in walking activities—either within-person or group activities—and over time) of walking behaviors [8,9].

Using this nuanced framework, we can consider how different attributes impact discrete walking behaviors. For example, we previously found that better physical function was associated with higher walking volumes in ARC residents in intermediate (ie, rest homes) and high-level (ie, hospitals) care, while surprisingly, moderate dementia, mild depression, and pain had no effect on walking volumes [10]. In contrast, people with mild cognitive impairment in the community show no differences in walking volumes compared to people who undergo normal aging, but they do demonstrate different patterns and greater variability in their walking behaviors [9]. By looking beyond the volume of walking activities to pattern and variability, we may garner information about people’s routines and the time they spend indoors and outdoors (based on walking bout lengths) [8,9,11-14]; we can then examine the impact of cognitive and physical impairments on these behaviors [14]. This information can contribute toward the development of more holistic interventions to support mobility in ARC.

Notably, ARC residents are a complex multimorbid population with significant variation in cognitive and physical function, often reflected in the level of care provided. These heterogeneities are not reflected in the literature regarding walking behaviors, highlighting a clear gap [8,15,16]. For example, Mc Ardle et al’s [8] review on the quantification of ambulatory activities in ARC reported that 26% of studies excluded people with cognitive impairment and only 17% explicitly characterized walking activities in people with cognitive impairment, despite a 65%-70% prevalence of cognitive impairment in ARC residents [17,18]. Additionally, no studies compared different levels of care. As such, Mc Ardle et al [8] recommended that we must characterize and compare the volumes, patterns, and variability of walking behaviors in ARC residents across different care levels, with different cognitive and physical abilities. By characterizing walking behaviors in a representative group of ambulatory ARC residents, we can gain a better understanding of physical mobility in ARC, which will inform future interventions and policies to promote walking activities and support mobility and function in ARC residents.

To address the highlighted gaps and recommendations, the primary aim of this study was to digitally characterize and compare walking behaviors across different levels of ARC using a validated and standardized framework, encompassing volume, pattern, and variability of walking. Secondary aims of this study were to characterize and compare walking behaviors in ARC residents according to their cognitive status and physical function.

Methods

Participants

Residents from 24 ARC facilities in New Zealand were recruited as part of the Staying UpRight randomized controlled trial (RCT), which evaluated an exercise intervention to reduce fall risk [19]. Only baseline data are included in this study. Participants were included if they were aged ≥65 years and mobile (ie, able to walk and transfer independently or with supervisory assistance). Participants were receiving one of the following levels of care: hospital-level care (24-hour care by, or under the supervision of, a nurse), rest home–level care (24-hour health-related care but not nursing care), or dementia-level care (rest-home level care in a secure environment to minimize the risks associated with dementia).

We excluded residents in psychogeriatric, respite, or palliative care; residents unable to undertake the assessment or the exercise intervention in the main RCT because they were acutely unwell (eg, gastroenteritis), or immobile (ie, unable to mobilize without 2-person assistance or bed bound) were also excluded.
Ethical Considerations
Participants who were able to give informed written consent did so before enrollment, and the facility clinical lead provided written consent for residents unable to provide their own informed consent because of cognitive impairment. The study was conducted according to the guidelines of the Declaration of Helsinki. Ethics approval was provided by the New Zealand Health and Disability Ethics Committee on October 31, 2018 (NZHDEC 18/NTB/151).

Clinical and Cognitive Outcomes
Demographic information for ARC residents included the following: age, sex, and years spent in the ARC facility. Physical function was measured using Short Physical Performance Battery (SPPB) [20] and the Timed Up and Go test [21]. Cognitive ability was assessed using the Montreal Cognitive Assessment (MoCA) [22].

Assessment of Walking Behaviors
ARC residents were asked to wear a small body-worn accelerometer (Axivity AX3; dimensions: 23 × 32.5 × 7.6 mm; weight: 11 g; sampling rate: 100 Hz; range: ±8 g; and memory: 512 MB) on the fifth lumber vertebra on the lower back. The accelerometer was affixed onto the skin using a double-sided hydrogel adhesive and a hypoallergenic plaster (Hypafix BSN Medical Limited). This particular protocol has been found to be feasible for multisite studies [23] in different aging cohorts [11,13,24]. Of particular note, algorithms used in this study for walking bout detection have been validated in ARC residents, with high accuracy for start and end time [25].

Participants were asked to wear the accelerometer continuously for 7 days, including in the shower and to bed. Once the assessment was complete, data were downloaded to a computer and processed via a validated analytical pipeline in MATLAB.

Data Processing and Walking Behavior Outcomes
Signals from the accelerometer were transformed to a horizontal-vertical co-ordinate system. Walking bouts were identified by filtering raw acceleration data using a second-order low-pass Butterworth two-pass digital filter, with a cutoff frequency of 17 Hz, and by applying selective thresholds on the vector magnitude and standard deviations of triaxial acceleration signals [11,23,26,27]. Once walking bouts were identified, for detecting steps, raw acceleration signals were filtered with low-pass, fourth-order Butterworth filter with cutoff frequency of 20 Hz. A Gaussian continuous wavelet transform of vertical acceleration was then applied to identify initial and final contacts, allowing the identification of steps. For each walking bout, total steps per bout and bout length were calculated. Sleep, lying, and sitting data were excluded based on the thresholds applied on the magnitude and standard of the accelerometry signal used to identify walking (eg, vertical acceleration, in a vertical position, needs to be −1 g and acceleration magnitude or standard needs to exceed these thresholds to be classified as walking). For sleep, the magnitude and standard of acceleration would be lower and the vertical acceleration would not be −1 g, so the position (orientation) excludes sleep, lying, or sitting.

A framework of walking behaviors was derived to remain consistent with previous literature [11,12,23,26], including volume, pattern, and variability of walking. Volume characteristics included total minutes spent walking as well as steps and bouts per day. Pattern characteristics included mean bout duration and alpha, which is derived by logarithmic transformation of bout density and length and is based on shape and power-law distribution [28,29]; alpha refers to the ratio of short to long walking bouts, which are scaled relative to an individual’s shortest walking bout. A high alpha score indicates that an individual’s total walking time is composed of proportionally shorter walking bouts compared to long walking bouts. Variability (S²) refers to the variability of bout duration between walking bouts, estimating how much an individual’s bout duration changes over the time period of data collection, and it was estimated using the maximum likelihood technique (previously described by Mc Ardle et al and Del Din [9,13]. The proportion of walking bouts taken in very short (<10 s), short (10-30 s), medium (30-60 s), and prolonged (>60 s) walks were calculated. These walking bout thresholds have been used commonly in other studies of a similar nature and provide contextual information regarding how walking takes place [13,30,31].

Considerations for Inclusion of Data
Given that most habitual walking takes place in <10-second bouts [13,32,33], we applied a minimum bout duration of 3 consecutive steps, and any period of rest that was ≥2.5 seconds was considered resting time [32]. Additionally, we included participants if they had ≥2 days of continuous walking activity data collected, as this is the minimum number of days required to reliably quantify our primary outcomes (ie, the volume of walking) across different care levels, based on Buckley et al [27].

Data Analysis
For demographic variables, chi-square tests and Fisher exact test were used to determine differences between groups for nominal variables, while one-way ANOVA was used to determine between-group differences for continuous variables; post hoc Tukey honest significant difference (HSD) tests determined where the differences lay.

Prior to statistical analysis relating to our primary and secondary aims, walking activity data were inspected visually using box plots, and outliers were identified. Separate analyses of covariance were used to assess differences in walking behaviors between groups categorized by level of care, cognition, or physical function while controlling for age and sex. Tukey HSD tests for multiple comparisons were used to determine where significant differences occurred. Sensitivity analysis was conducted by removing outliers more than 1.5 times above the third quartile or below the first quartile and by conducting the analysis of covariance and subsequent post hoc tests for each discrete grouping separately (eg, level of care, cognition, or physical function).

The effect size of group differences was calculated using the Hedges g formula to account for disparities between groups’ sample sizes [34]. Effect sizes are interpreted as follows: 0.2-0.4: ...
small, 0.5-0.7: medium, and ≥0.8: large. Assumptions were evaluated (eg, normality of residuals) for all models, and statistical significance was defined as a $P<.05$.

Cognitive levels were assessed and categorized using MoCA cutoff scores, as follows: cognitively intact (MoCA ≥26), mild cognitive impairment (MoCA 18-25), moderate cognitive impairment (MoCA 10-17), and severe cognitive impairment (MoCA <10) [35]. Cognitively intact participants were excluded from the cognitive impairment severity analysis due to the small sample size but retained for illustrative purposes in Figures. Physical function levels were assessed and categorized using the SPPB cutoff scores, as follows: high-moderate function (SPPB 12-7) or low-very low function (SPPB <7) [36].

Results

Demographic Information

A total of 306 ARC residents were included in this analysis and were primarily grouped according to their level of care (Table 1). Figure 1 outlines reasons for exclusion and inclusion of participants for this secondary analysis from the Staying UpRight RCT. Hospital-level care residents had lower physical function compared to rest home–level care residents ($P=.01$) and took a longer time to complete the Timed Up and Go test compared to rest home–level ($P<.001$) and dementia-level ($P=.03$) residents. MoCA scores were significantly lower in dementia-level residents compared to rest home–level and hospital-level care residents ($P<.001$).
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Hospital (n=117)</th>
<th>Rest home (n=164)</th>
<th>Dementia care (n=25)</th>
<th>Overall P value&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years; n=306), mean (SD)</strong></td>
<td>84 (7)</td>
<td>84 (7)</td>
<td>81 (8)</td>
<td>.20</td>
</tr>
<tr>
<td><strong>Sex (n=306), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>70 (60)</td>
<td>101 (62)</td>
<td>16 (64)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>47 (40)</td>
<td>63 (38)</td>
<td>9 (36)</td>
<td>.90</td>
</tr>
<tr>
<td><strong>Years in facility&lt;sup&gt;b&lt;/sup&gt; (n=304), mean (SD)</strong></td>
<td>0.4 (0.2)</td>
<td>0.4 (0.2)</td>
<td>0.4 (0.1)</td>
<td>.60</td>
</tr>
<tr>
<td><strong>SPPB&lt;sup&gt;c&lt;/sup&gt; score (0-12; n=296), mean (SD)</strong></td>
<td>4.3 (2.6)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>5.2 (2.6)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>4.1 (2.3)</td>
<td>.008</td>
</tr>
<tr>
<td><strong>Physical function level&lt;sup&gt;e&lt;/sup&gt; (n=296), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High physical function (SPPB 10-12)</td>
<td>5 (4.3)</td>
<td>9 (5.6)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Moderate physical function (SPPB 7-9)</td>
<td>19 (17)</td>
<td>38 (24)</td>
<td>3 (15)</td>
<td></td>
</tr>
<tr>
<td>Low physical function (SPPB 4-6)</td>
<td>37 (32)</td>
<td>70 (43)</td>
<td>8 (40)</td>
<td></td>
</tr>
<tr>
<td>Very low physical function (SPPB &lt;4)</td>
<td>54 (47)</td>
<td>44 (27)</td>
<td>9 (45)</td>
<td></td>
</tr>
<tr>
<td>Unknown&lt;sup&gt;f&lt;/sup&gt;</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td><strong>TUG&lt;sup&gt;h&lt;/sup&gt; (s; n=289), mean (SD)</strong></td>
<td>37 (22)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>27 (15)&lt;sup&gt;d&lt;/sup&gt;</td>
<td>25 (18)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>MoCA&lt;sup&gt;j&lt;/sup&gt; score (0-30; n=285), mean (SD)</strong></td>
<td>15 (6)&lt;sup&gt;i&lt;/sup&gt;</td>
<td>15 (6)&lt;sup&gt;k&lt;/sup&gt;</td>
<td>4 (6)&lt;sup&gt;1&lt;/sup&gt;</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Cognitive level&lt;sup&gt;e&lt;/sup&gt; (n=285), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognitively intact (MoCA ≥26)</td>
<td>5 (4.6)</td>
<td>4 (2.5)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Mild cognitive impairment (MoCA 18-25)</td>
<td>28 (26)</td>
<td>58 (36)</td>
<td>1 (5.6)</td>
<td></td>
</tr>
<tr>
<td>Moderate cognitive impairment (MoCA 10-17)</td>
<td>60 (56)</td>
<td>66 (42)</td>
<td>2 (11)</td>
<td></td>
</tr>
<tr>
<td>Severe cognitive impairment (MoCA &lt;10)</td>
<td>15 (14)</td>
<td>31 (19)</td>
<td>15 (83)</td>
<td></td>
</tr>
<tr>
<td>Not tested&lt;sup&gt;d&lt;/sup&gt;</td>
<td>9</td>
<td>5</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td><strong>Days wearing the activity monitor (n=306), mean (SD)</strong></td>
<td>6.5 (1)</td>
<td>6.4 (1)</td>
<td>6.3 (1)</td>
<td>.60</td>
</tr>
</tbody>
</table>

<sup>a</sup>One-way ANOVA, Pearson chi-square test, and Fisher exact test.

<sup>b</sup>For years in facility, 1 participant’s data were missing from both the “hospital” and “rest home” groups.

<sup>c</sup>SPPB: Short Physical Performance Battery (2 participants in the “hospital” group, 3 participants in the “rest home” group, and 5 participants in the “dementia care” group were not tested for SPPB).

<sup>d</sup>Hospital vs rest home.

<sup>e</sup>Descriptive variable only (no statistical testing performed).

<sup>f</sup>Not applicable.

<sup>g</sup>“Unknown” indicates participant data missing in each group, so percentages are not applicable.

<sup>h</sup>TUG: Timed Up and Go (6 participants in the “hospital” group, 3 participants in the “rest home” group, and 8 participants in the “dementia care” group were not tested).

<sup>i</sup>Hospital vs dementia care.

<sup>j</sup>MoCA: Montreal Cognitive Assessment (9 participants in the “hospital” group, 5 participants in the “rest home” group, and 7 participants in the “dementia care” group were not tested).

<sup>k</sup>Rest home vs dementia care.

<sup>l</sup>Indicates the number of participants not tested in each group, so percentages are not applicable.
Walking Behaviors Across Care Levels

Dementia care residents demonstrated higher volumes of walking, with longer, more variable bout durations and lower alpha scores compared to residents in both rest homes and hospitals (moderate to large effect sizes). Residents in rest homes also showed higher volumes of walking compared to those in hospitals (small effect sizes). Table 2 provides further details; Multimedia Appendix 1 provides P values for post hoc tests and details regarding effect sizes. Notably, sensitivity analysis indicated that variability of walking bout length did not differ between groups following the removal of outliers (P= .16).
Characterization of walking behaviors, categorized by level of care (N=306).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Hospital (n=117)</th>
<th>Rest home (n=164)</th>
<th>Dementia care (n=25)</th>
<th>Overall P valuea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walk time per day (min), mean (SD)</td>
<td>58 (37)b,c</td>
<td>74 (39)b,d</td>
<td>137 (59)c,d</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Steps per day, mean (SD)</td>
<td>4138 (2766)b,c</td>
<td>5216 (2925)b,d</td>
<td>10,886 (5453)c,d</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Bouts per day, mean (SD)</td>
<td>256 (165)b,c</td>
<td>321 (160)b,d</td>
<td>496 (238)c,d</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Mean bout duration (s), mean (SD)</td>
<td>13.9 (3.6)c</td>
<td>14.1 (3.3)d</td>
<td>20.1 (20.4)c,d</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Variability, mean (SD)</td>
<td>0.81 (0.11)c</td>
<td>0.80 (0.09)d</td>
<td>0.89 (0.18)c,d</td>
<td>.002</td>
</tr>
<tr>
<td>Alpha score, mean (SD)</td>
<td>1.68 (0.08)c</td>
<td>1.67 (0.07)d</td>
<td>1.61 (0.09)c,d</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Distribution of walking bouts by discrete walking bout length (%), mean (SD)

<table>
<thead>
<tr>
<th></th>
<th>Hospital (n=117)</th>
<th>Rest home (n=164)</th>
<th>Dementia care (n=25)</th>
<th>Overall P valuea</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;10-second bouts</td>
<td>65 (8)c</td>
<td>64 (8)d</td>
<td>56 (10)c,d</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>10- to 30-second bouts</td>
<td>25.9 (5.5)c</td>
<td>27.6 (5.5)d</td>
<td>30.8 (5.8)c,d</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>30- to 60-second bouts</td>
<td>5.81 (3.05)c</td>
<td>5.51 (2.42)d</td>
<td>8.64 (3.64)c,d</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>&gt;60-second bouts</td>
<td>3.20 (2.13)c</td>
<td>3.15 (2.06)d</td>
<td>4.82 (4.92)c,d</td>
<td>.01</td>
</tr>
</tbody>
</table>

*One-way ANOVA, controlling for age and sex.

bHospital vs rest home.

cHospital vs dementia care.

dRest home vs dementia care.

Additionally, dementia care residents spent a significantly lower percentage of their walking bouts in very short bouts (eg, <10 s) and a greater percentage in short, medium, and prolonged walking bouts compared to residents in other care levels (moderate to large effect sizes; Table 2).

Walking Behaviors Across Cognitive Impairment Severities

There were no significant differences between cognitive groups for any volume characteristics (Figure 2 and Multimedia Appendix 2). People with severe cognitive impairment took longer, more variable walking bouts with a lower alpha score compared to those with mild (moderate to large effect sizes) and moderate cognitive impairment (small to moderate effect sizes). Figure 2 and Multimedia Appendix 2 provide further details.
Walking Behaviors Across Physical Function Levels
ARC residents with high-moderate physical function spent more time walking and took more steps and bouts per day, with less variability for bout length, compared to those with low-very low physical function (Figure 3 and Multimedia Appendix 3 present further details).
Discussion

Principal Findings
This is the first study to describe the volume, pattern, and variability of walking behaviors in ARC, captured by an accelerometer, with consideration of different care levels, cognitive abilities, and physical function, reflecting the typical population of residents. Key findings highlight that dementia care residents have significantly higher volumes of walking, take longer and more variable walking bouts on average, and spend proportionately more of their walking time in prolonged bouts of walking compared to rest home and hospital levels of care. Although the volume of walking is similar across different levels of cognitive impairment severity, people with more severe cognitive impairment show different patterns (eg, longer walking bouts) and greater variability compared to those who are less cognitively impaired. In contrast, people with lower physical function have significantly lower volumes of walking and higher variability of walking bout lengths but do not differ in terms of the pattern of this activity. ARC residents are a complex multimorbid population who often present with varying levels of both cognitive and physical abilities, and these nuances should be considered in further research aiming to improve mobility and reduce fall risk.
Walking Behaviors Across Different Care Environments

This is the first study to show that people living in a dementia unit participate in higher volumes of walking, with different patterns (ie, longer walking bouts) and greater variability compared to other ARC environments, with medium to large effect sizes. However, differences between groups for the variability of walking bout length disappeared following the removal of significant outliers, therefore, results should be interpreted with caution. As physical function scores are comparable between the dementia care and rest home residents (Table 1), the differences in walking behaviors may illustrate a behavioral component of dementia (eg, wandering—a dementia-related locomotor behavior involving frequent and repetitive movements, such as pacing). Objective remote monitoring of wandering behaviors using digital methods has previously been proposed to detect and monitor wandering behaviors [37,38]. We propose that variability of walking bout lengths should be considered in future research in this area, as it may reflect wandering behaviors [39]; clinical validation is required to investigate this hypothesis.

Contrary to our findings, baseline results from one previous RCT reported very low volumes of walking activity in dementia care units, showing <400 steps in one 24-hour period assessed via an activity armband [40]. However, these studies are difficult to compare due to differences in device location (ie, lower back compared to arm) and data collection periods. Moyle et al [40] noted that this activity armbands were unreliable and resulted in large amounts of missing data. Additionally, high volumes of walking activity reported here (eg, >4000 steps per day) are likely due to our low cutoff thresholds for defining walking activity, as most walking takes place in very short walking bouts in this population. The cutoff threshold applied to characterize walking behaviors can make significant differences in the volume of walking captured—with differences ranging from 2000 to 10,000 steps in previous literature [30]. This study has expanded beyond simple volume metrics and highlighted the importance of selecting validated and sensitive digital methods when assessing walking behaviors in this population [8].

Walking Behavior Across Cognitive Impairment Severity and Physical Capacities

This is the first study to demonstrate that people with severe cognitive impairment have similar volumes but significantly different patterns (ie, longer bouts) and greater variability of walking compared to less cognitively impaired groups, suggesting that although cognition does not influence the amount of activity, it may change the way this activity is carried out. This finding is supported by and extends our previous work, which excluded dementia care residents and highlighted that while worse physical function is associated with lower volumes of walking in ARC, cognitive impairment showed no effect on walking volume [10]. Perhaps this indicates that pattern and variability of walking behaviors are cognitively mediated outcomes and may be useful to monitor in ARC as a proxy for cognitive decline. For example, in line with our results from the dementia care unit, the literature indicates that people with severe dementia are more likely to wander [41] and we propose that this is reflected in the pattern and variability of walking. Longer walking bouts and higher variability of bout length are considered positive outcomes in cognitively healthy individuals, indicative of dynamic and varied routines [11], but perhaps higher variability in tandem with significant cognitive impairment is more reflective of repetitive lapping behaviors (ie, wandering). Clinical validation is required to address this speculation. Differences in pattern of walking behaviors have previously been reported between community-dwellers with mild cognitive impairment and normal aging [42], supporting the hypothesis that cognitive decline may influence these behaviors.

In contrast, people with worse physical function have significantly lower volumes of walking but show no differences in pattern or variability compared to those with better physical function. The association between higher walking volumes and better physical function confirms the findings of previous studies [43,44]. However, the cross-sectional design of this study precludes commentary on the direction of causality. Although from our results, we cannot determine if encouraging walking as part of a resident’s daily activities can result in clinically meaningful improvements in function, previous research demonstrated that function-focused care (ie, increasing routine activities) leads to increased activity volumes and improved functional outcomes in ARC residents with moderate functional dependency [44] but not in dementia residents with severe functional dependency [45]. However, pattern and variability of walking are considered to reflect daily routines, and the effects of function-focused care may be more readily observed in these outcomes rather than in volume, especially in individuals with severe cognitive impairment. Additionally, marginal increases in the duration and variability of walking bouts may lead to significant improvements in function [28] and should be considered in very frail residents. These hypotheses could be considered in future intervention studies, with consideration for the multimorbid nature and varying levels of both cognitive and physical issues inherent in ARC residents.

Strengths and Limitations

Strengths of this study were the large sample, drawn from multiple facilities, distributed across 3 levels of care and encompassing a broad spectrum of cognitive and physical capacities. This is particularly notable, as there can be significant difficulties in collecting data using wearable technology from people with dementia in ARC facilities [29]. We used a technically appropriate digital method to collect low volumes of walking data, meeting the recommendations from Mc Ardle et al [8]. Additionally, we used a standardized framework to characterize walking behaviors, making our findings comparable to multiple other cohorts and enhancing our understanding of walking behaviors across the spectrum of care and cognition [11,12]. We addressed the reliability of our primary outcomes based on previous empirical evidence [27].

Our study has several limitations. Residents were only included if they could ambulate, and residents who could not complete the MoCA or SPPB were not included in our secondary analysis; therefore, we may have reduced representation of different levels of cognitive and physical capacities. Although this is only the
second study to specifically characterize walking activity in a dementia care unit [8], it should be noted that our sample size for this group was low and likely to have limited statistical power; therefore, statistical analysis was exploratory and results should be considered with that in mind. This is a cross-sectional study; therefore, assessing changes in walking activity over time or establishing causality of influences on walking activity is not possible; in the future, a longitudinal study may offer valuable insights into predictors of walking behaviors in ARC. Although we adjusted for multiple corrections within statistical models (ie, Tukey HSD tests), we did not adjust for multiple comparisons for multiple outcomes, and there may be a risk of type I error. We also included participants with ≥2 days of walking activity data, as this is the number of days required to obtain reliable volume outcomes (ie, our primary outcome) in ARC [27]; however, our secondary outcomes of pattern and variability require 2-5 days of data to ensure reliability, pending on the discrete variable, and thus, results should be interpreted with caution. Although commonly assessed in ARC [8], we did not include outcomes relating to the intensity of walking activity, as this has been suggested to be inappropriate to characterize in this population, given that ARC residents primarily engage only in light-intensity activities [8]. Additionally, the ARC facilities included in this study reflect a New Zealand context, and findings may be different in other countries due to alternative organizational features and policies [46]. We recognize that apart from resident-related factors of physical function and cognition, walking activity may be influenced by the physical and organizational environment [46]—aspects that were not measured in our study. As previously noted, ARC residents may have varying levels of both cognitive and physical impairments, and the combined spectrum should be considered in future research. Finally, digital outcomes beyond those described in this analysis can provide important clinical information about ARC residents and should be considered in future research. For example, sleep disturbances can be measured using actigraphy. Sleep disturbances are common in people living in ARC and are associated with neuropsychiatric symptoms and prescription of psychotropic drugs, which can enhance fall risks and greater staff distress [47]. Although it is beyond the scope of this study, further research may consider using qualitative approaches to complement current findings and the wider literature [8], which would allow us to garner rich insights from ARC residents regarding which digital outcomes relate to their lived experiences and are meaningful to assess.

Conclusions
This is the first study to show the influence of care environment, cognitive status, and physical function on walking behaviors in ARC residents. Our results indicate that cognitive and physical abilities may discretely impact the volumes, pattern, and variability of walking. This work has addressed a significant gap in the literature and has generated new hypotheses regarding which digitally derived walking outcomes are meaningful to assess in ARC residents.

Acknowledgments
The study has been funded by a Health Research Council of New Zealand project grant (reference 18/414). The study funder plays no role in study design, data collection, or analyses.

RMA received funding from the National Institute for Health and Care Research (NIHR) for her fellowship (NIHR 301677). SDD and LR were supported by the Innovative Medicines Initiative 2 Joint Undertaking (IMI2 JU) project IDEA-FAST (grant agreement 853981). This joint undertaking (JU) receives support from the European Union’s Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations (EFPIA). SDD and LR were also supported by the Mobilise-D project, which has received funding from the Innovative Medicines Initiative 2 JU (grant agreement 820820). RMA, SDD, and LR were supported by the NIHR Newcastle Biomedical Research Centre (BRC) based at The Newcastle Upon Tyne Hospital National Health Service (NHS) Foundation Trust, Newcastle University, and the Cumbria, Northumberland and Tyne and Wear (CNTW) NHS Foundation Trust. SDD and LR were supported by the NIHR/Wellcome Trust Clinical Research Facility (CRF) infrastructure at Newcastle Upon Tyne Hospitals NHS Foundation Trust.

All opinions are those of the authors and not the funders.

The authors would like to acknowledge the care home residents who participated in this study and the study research assistants for their time in data collection and analysis.

Authors’ Contributions
RMA, LT, LR, SDD, and NK contributed to the conceptualization and design of the study as well as data analysis and interpretation. RMA, LT, and SDD were in charge of data curation. RMA conducted data analysis and statistical work. AC provided support for statistical work. RMA and LT wrote the original draft. RMA, LT, LR, AC, SDD, and NK wrote the draft, reviewed, and edited it. All authors approved the final version of this manuscript. RMA and LT contributed equally to this manuscript; SDD and NK also contributed equally to this manuscript.

Conflicts of Interest
SDD reports consultancy activity with Hoffmann-La Roche Ltd. LR consults for MJ Fox Foundation for service on Endpoints Advisory Committee.
Multimedia Appendix 1
Detailed description of between-group analysis results for different care levels.
[DOCX File, 13 KB - aging_v7i1e53020_app1.docx]

Multimedia Appendix 2
Walking behaviors categorized by cognitive impairment severity.
[DOCX File, 19 KB - aging_v7i1e53020_app2.docx]

Multimedia Appendix 3
Walking behaviors categorized by physical function.
[DOCX File, 17 KB - aging_v7i1e53020_app3.docx]

References


Abbreviations

ARC: aged residential care
HSD: honest significant difference
MoCA: Montreal Cognitive Assessment
RCT: randomized controlled trial
SPPB: Short Physical Performance Battery
**Digital Peer-Supported App Intervention to Promote Physical Activity Among Community-Dwelling Older Adults: Nonrandomized Controlled Trial**

Kento Tabira\(^1,2\), MSc, PT; Yuko Oguma\(^1,2\), MPH, MD, PhD; Shota Yoshihara\(^3,4\), MSc, OT; Megumi Shibuya\(^4\), MSc; Manabu Nakamura\(^1,2\), MSc, PT; Natsue Doihara\(^2\), PhD; Akihiro Hirata\(^1,2,5\), MSc; Tomoki Manabe\(^1,2\), RN, MPH, PHN

\(^1\)Graduate School of Health Management, Keio University, Kanagawa, Japan  
\(^2\)Sports Medicine Research Center, Keio University, Kanagawa, Japan  
\(^3\)Department of Rehabilitation Sciences, Graduate School of Medical Sciences, Kitasato University, Kanagawa, Japan  
\(^4\)A10 Lab Inc, Tokyo, Japan  
\(^5\)Japan Society for the Promotion of Science, Tokyo, Japan

**Corresponding Author:**  
Yuko Oguma, MPH, MD, PhD  
Sports Medicine Research Center, Keio University  
4-1-1 Hiyoshi, Kohoku-ku, Yokohama  
Kanagawa, 223-0061  
Japan  
Phone: 81 45 566 1090  
Email: yoguma@keio.jp

**Abstract**

**Background:** The use of mobile apps has promoted physical activity levels. Recently, with an increasing number of older adults accessing the internet, app-based interventions may be feasible in older populations. Peer support–based interventions have become a common method for promoting health-related behavior change. To our knowledge, the feasibility of using digital peer support apps (DPSAs) to increase physical activity among older adults and its impact on physical activity and physical function have not been investigated.

**Objective:** This study aims to assess the feasibility of using DPSAs in older adults and to assess changes in physical activity and physical function in DPSA users.

**Methods:** We conducted a nonrandomized controlled trial of older adults aged ≥65 years. We recruited participants for 2 distinct 12-week programs designed to increase physical activity. Participants could choose between an intervention group (app program and exercise instruction) or a control group (exercise instruction only). DPSA creates a group chat for up to 5 people with a common goal, and participants anonymously post to each other in the group. Once a day, participants posted a set of their step counts, photos, and comments on a group chat box. The intervention group used the DPSA after receiving 2 face-to-face lectures on its use. The participants were characterized using questionnaires, accelerometers, and physical function assessments. The feasibility of the DPSA was assessed using retention and adherence rates. Physical activity was assessed using accelerometers to measure the daily step count, light intensity physical activity, moderate to vigorous intensity physical activity (MVPA), and sedentary behavior. Physical function was assessed using grip strength and the 30-second chair-stand test.

**Results:** The participants in the intervention group were more frequent users of apps, were more familiar with information and communication technology, and had a higher baseline physical activity level. The retention and adherence rates for the DPSA intervention were 88% (36/41) and 87.7%, respectively, indicating good feasibility. Participants in the intervention group increased their step count by at least 1000 steps and their MVPA by at least 10 minutes using the DPSA. There was a significant difference in the interaction between groups and intervention time points in the daily step count and MVPA (step count, \(P=.04\); duration of MVPA, \(P=.02\)). The DPSA increased physical activity, especially in older adults with low baseline physical activity levels.

**Conclusions:** The feasibility of DPSA was found to be good, with the intervention group showing increases in daily steps and MVPA. The effects of DPSA on step count, physical activity, and physical function in older adults with low baseline physical activity should be investigated using randomized controlled trials.

https://aging.jmir.org/2024/1/e56184
Introduction

Background
The health benefits of regular physical activity are familiar [1]. Physical activity reduces the risk of chronic diseases such as type 2 diabetes, cardiovascular disease, and hypertension [2-4]. In addition, physical activity improves the overall physical and mental functioning and controls morbidity and mortality rates [5,6]. However, globally, the level of physical activity has remained stable or declined, despite several efforts to promote physical activity [7,8]. In addition, social distancing during the COVID-19 pandemic has caused changes in lifestyle and social behavior [9]. The level of physical activity among older adults in Japan is reported to have declined due to the COVID-19 pandemic [10,11] and needs to be increased.

Recently, mobile apps have been used successfully to increase physical activity levels [12,13]. eHealth encompasses health care services and information delivered with the aid of information and communication technology (ICT), including computers, mobile phones, and satellite communications. Mobile health (mHealth) refers to the use of smart or portable devices for providing health services and information [14]. With an increase in the population of older adults using the internet, mHealth and eHealth approaches may be feasible [15,16]. A total of 3 out of 4 reviews concluded that mHealth or eHealth interventions are effective in the short term in promoting physical activity in adults aged ≥50 years [17]. Furthermore, eHealth interventions targeting physical activity have revealed that theory-based interventions are more effective than interventions that are not grounded in theory [18]. However, there has been limited focus on social cognitive theory–based interventions aimed at promoting physical activity among older adults through peer support.

Peer support–based interventions have become a common method for promoting health-related behavior change [19]. Webel et al [20] defined peer support as “a method of teaching or facilitating health promotion that makes use of people sharing specific health messages with members of their own community.” The effectiveness of peer support–based interventions for physical activity has a theoretical basis, often explained by social cognitive theory [21]. The social cognitive theory proposed by Bandura [22] stipulates that behavior is learned by observing and imitating others. This process is called observational learning or modeling and has been extensively studied in the context of motor skill development and education. Peer-mediated delivery of information regarding physical activity through apps could facilitate attention, retention, and motivation to work on that information, as per social cognitive theory. Liu and Lachman [23] conducted a 4-week randomized controlled trial based on social cognitive theory in which older adults aged ≥60 years used the WeChat and WeRun apps to increase their step counts by recording and exchanging them through SMS text messages. This social engagement through SMS text messages increased the step count. However, the step count was the only physical activity outcome measure, and the effect of the intervention on physical activity intensity, sedentary behavior (SB), and physical function was not assessed. Therefore, using an app based on social cognitive theory, we examined the effects of a digital peer-supported intervention on step counts, physical activity intensity, SB, and physical function among older adults aged ≥65 years.

Objective
This study used a digital peer-supported app (DPSA) to conduct a 12-week intervention study on older adults aged ≥65 years. The objectives of this study were threefold: (1) to understand the characteristics of older adults who choose to use the DPSA to increase their physical activity; (2) to evaluate the feasibility of using the DPSA to promote physical activity in older adults; and (3) to measure the effect of using the DPSA on users’ level of physical activity, SB, physical function, and self-efficacy for exercise.

Methods

Study Design
This nonrandomized pretest-posttest comparison trial of 2 groups was conducted in Fujisawa City, Kanagawa Prefecture, Japan. In April 2022, the city had an area of 69.57 km² and a population of 442,892, of whom 108,472 (24.49%) were aged ≥65 years. The study was conducted as an industry-government-academia collaboration between the local government, an app-making company, and a university.

Ethical Considerations
The study was approved by the Research Ethics Committee of Sports Medicine Research Center at Keio University (approval number 2022-07). Informed consent was obtained from all participants. The study protocol was registered with the University Hospital Medical Information Network (UMIN000050618).

Participants
The study was conducted on Fujisawa City older adults aged ≥65 years. We recruited participants for 2 distinct 12-week programs designed to increase physical activity. Participants could choose between an intervention group (app program and exercise instruction) and a control group (exercise instruction only). Participants from different areas within the municipality were recruited through flyers, publicity, and calls to related organizations. The intervention was implemented in two phases: (1) from October 2022 to January 2023 and (2) from December 2022 to March 2023. Participants selected programs according
to their preferences (the app program was for smartphone owners only). The eligibility criteria were older adults aged ≥65 years, who were able to walk independently and perform other activities of daily living and had not been advised to refrain from physical activity by a physician. Before participation, prospective participants were screened using a personal health status questionnaire based on the Physical Activity Readiness Questionnaire [24-26] to ascertain whether there were any potential problems with participation in the study. Because older adults are generally assumed to be less familiar with smartphones and apps than younger adults [27] and may not be able to completely use the DPSA, we provided 2 lectures about using the app; 1 lecture was conducted at the start of the DPSA use, and another was conducted 1 week later. Participants were instructed to download the app, and its use was explained during the first lecture. Each lecture lasted for 1 to 1.5 hours, and the participants were able to receive instructions directly from the instructor and ask questions.

**Intervention**

**Program**

The timeline of the study procedure is presented in Figure 1. Regardless of program selection, all participants participated in face-to-face exercise instruction, program introduction, and baseline assessment conducted by a physical therapist or a health fitness instructor. Exercise instructions focused on aerobic, stretching, muscle strengthening, and balance exercises based on the original “Fujisawa +10 exercise” program [28,29]. Exercise instruction was provided for 15 minutes for the intervention group and 30 minutes for the control group. Both intervention and control groups were instructed to increase their daily physical activity. Participants completed questionnaires and underwent physical activity and physical function assessments using a triaxial accelerometer at baseline (start date) and during follow-up (weeks 10-12). Individual physical activity reports were generated from the data obtained and fed back to the participants. The intervention group began using the app 1 week after the baseline outcome assessment was conducted. The timing of the evaluation of each measurement item is described in Multimedia Appendix 1.

**Figure 1.** Timeline of study procedures. DSPA: digital peer support app; wk: weeks.

<table>
<thead>
<tr>
<th>Start Date</th>
<th>2 wk</th>
<th>4 wk</th>
<th>6 wk</th>
<th>8 wk</th>
<th>10 wk</th>
<th>12 wk</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follow-up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exercise instruction ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assessment (Except physical activity) ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Physical activity assessment 1 wk ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1 wk</td>
<td></td>
</tr>
<tr>
<td>Physical activity report distribution ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Use of DPSA (Intervention group only) ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>12 wk</td>
<td></td>
</tr>
<tr>
<td>Lecture on app use (Intervention group only) ✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>

**DPSA Intervention**

The study was conducted using Minchalle (A10 Lab Inc), a commercially available DPSA [30]. This app was developed in June 2015 and launched in November 2015. Figure 2 shows example screens from the app. The DPSA creates a group chat for up to 5 people with a common goal, and participants anonymously post messages to each other in the group. The common goal of the intervention group was to increase physical activity through walking and exercise. Once a day, participants posted a set of their step counts, photos, and comments on a group chat box. The main functions of the DPSA used in this study were to enable the participants to (1) post step counts, photos, and comments about the day; (2) post approvals from group members to each other’s postings; (3) set step count goals on a group basis; and (4) get feedback on the team’s total daily step count. Step counts were measured using a smartphone, with the DPSA reporting the number of steps taken on the day at the time of posting. Participants had the option to post comments or photos multiple times a day and engage with other members, although this was not mandatory. The app was available for participants to use free of charge.
Figure 2. Examples of app screens. (A) Select a group. (B) Post photos, step count, and comments on the group. Post a photo taken that day and comment on the day’s events. (C) The contents of the posts are displayed in the group. The total number of steps for the group is displayed. (D) Response to posts by group members.

Measurements

Demographic Characteristics of Participants

In addition to general characteristics such as age and sex, the survey asked about smartphone ownership, the frequency of app use, exercise habits, the frequency of neighborhood interaction, participation in group exercise, history of falls in the past year, the effect of the COVID-19 pandemic on their level of physical activity, self-reported decrease in walking speed, the use of ICT, and self-efficacy for exercise.

Participant’s body weight (kg) was measured using a digital scale, and height (m) was measured on a stadiometer after participants removed their shoes. BMI was then calculated as body weight divided by the square of height.

Exercise habits were defined as “those who exercise at least twice a week, for at least 30 minutes each time, for at least one year” [31].

The frequency of neighborhood interaction was assessed by asking how many times one interacts with people in the neighborhood within 1 week.

Group exercise participation was defined as those who participate in a group of three or more people who meet voluntarily to exercise.

Information about the use of ICT was collected using questions about “Gathering information and shopping on the internet,” “Using social networking services (Facebook, LINE, Instagram, etc),” and “Do not use any information devices.” The percentages for the intervention and control groups were compared with representative values from the Annual Report on the Ageing Society, published by the Japanese Cabinet Office, to determine the extent to which participants are using ICT compared with other older adults [32].

Outcome Measures of Participants

To assess physical activity, participants were asked to wear a triaxial accelerometer [33] (Active Style Pro HJA-750C Activity Meter, Omron Health Care) at waist level for 7 consecutive days before the intervention and 10 to 11 weeks after the intervention commenced. The accelerometer display was configured to prevent users from viewing the amount of physical activity for the day. Participants were instructed not to remove the device unless required for certain tasks, such as changing their clothes and bathing. At the end of the measurement, all the data collected were transferred from the accelerometer to a PC. Following the method suggested by Jefferis et al [34] for estimating physical activity, an individual needed to record at least 10 hours of activity per day for 3 days to be included in the subsequent analyses. The data were collected in 60-second epochs for data analysis and used to estimate the intensity of activity in metabolic equivalents (METs). The mean daily step count and time spent in SB (≤1.5 METs), light intensity physical activity (LPA; 1.6-2.9 METs), and moderate to vigorous intensity physical activity (MVPA: ≥3 METs) per day were used for outcome measurements of physical activity.

Physical function was assessed using grip strength and the 30-second chair-stand test (CS-30). Grip strength was measured using a digital dynamometer (Grip D, TKK 5401, Takei Scientific Instruments). Measurements were taken in the standing position, with the elbow joint in extension and the wrist joint in midextension. Both the left and right hands were measured once each, and the maximum value was used. For the CS-30 test [35], seated participants were instructed to stand up from the chair with their arms crossed at chest level as many times as possible in 30 seconds.

Self-efficacy for exercise was assessed using 4 questions pertaining to participants’ self-confidence in exercising under each of the following conditions [36]: “Do you have the confidence to exercise regularly under the following conditions? physical fatigue, mental stress, lack of time, and bad weather.”
In response to the question, participants were asked to select 1 of the 5 answers ranging from “No, I don’t have any confidence at all (1 point)” to “Yes, I am quite confident (5 points).” The

Textbox 1. Measurements related to participant characteristics.

<table>
<thead>
<tr>
<th>Measurement methods and items</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Questionnaire</td>
</tr>
<tr>
<td>• Age, sex, living alone, self-rated health, perceived household economic status, life satisfaction, employment status, smartphone ownership, frequency of app use, exercise habits, frequency of neighborhood interaction, participation in group exercise, history of falls in the past year, effect of the COVID-19 pandemic on decreased physical activity, self-reported decrease in walking speed, and self-efficacy for exercise</td>
</tr>
<tr>
<td>• Triaxial accelerometer</td>
</tr>
<tr>
<td>• Steps, light intensity physical activity, moderate to vigorous intensity physical activity, and sedentary behavior</td>
</tr>
<tr>
<td>• Physical function assessment</td>
</tr>
<tr>
<td>• BMI, grip strength, and 30-second chair-stand test</td>
</tr>
</tbody>
</table>

**Feasibility of DPSA Intervention**

The feasibility of DPSA intervention was assessed by retention and adherence rates during the 12-week program implementation. The DPSA could exclude a person from a group if they have not posted a set of their step counts, photos, and comments for 15 consecutive days. Dropouts were defined as those who were excluded from the group during the 12 weeks of DPSA intervention. The retention rate was calculated using a denominator of 41 participants including those who withdrew consent. The adherence rate of DPSA intervention was calculated by dividing the number of sets of their step counts, photos, and comments posted during the intervention period by the duration of the intervention. DPSA adherence rates were also calculated by group (9 groups: A-I). The number of all chat posts per person by group was calculated to assess group use. Negative physical effects that occurred during the intervention were ascertained by interviewing the participants during follow-up. We report on privacy breaches and technical problems with the app. Privacy breaches were identified by the municipality, and technical problems were identified by the app company. Continuity was evaluated using a questionnaire on factors that contributed to exercise continuation by the DPSA and the intention to continue using the DPSA after 12 weeks.

**Changes in Physical Activity, SB, Physical Function, and Self-Efficacy for Exercise**

Physical activity (step count, LPA, MVPA, and SB); physical function (grip strength and CS-30); and self-efficacy for exercise were measured at 2 time points: baseline and follow-up. The follow-up data were measured in the same way as at baseline. For follow-up data, physical activity was measured between weeks 10 and 11 of the intervention. Physical function and self-efficacy for exercise were measured after 12 weeks of the intervention.

**Statistical Analysis**

Comparisons of participant characteristics in each group were analyzed using the independent samples $t$ test (2-tailed), chi-square test, and Mann-Whitney $U$ test. The interaction between the group and time of intervention was analyzed using a linear mixed model with baseline and follow-up group differences adjusted for baseline age, sex, and frequency of app use (at baseline). Consent withdrawers were excluded, and older adults who dropped out of the intervention were included in the analysis in a modified intention-to-treat analysis. The daily step count, SB, LPA, MVPA, grip strength, CS-30, and self-efficacy for exercise were analyzed as dependent variables in separate models. The daily step counts were positively skewed; therefore, square root transformations were applied to improve normality. To increase the comprehensibility of the tables, raw descriptive data were reported, although analyses were conducted using the square root–transformed values. We defined high and low levels of physical activity based on step counts, using 7000 steps per day as the cutoff, as the recommended step goal for older adults is typically 7000 to 10,000 steps [37]. To examine the impact of physical activity at baseline on the intervention effect, a post hoc subgroup analysis was conducted on participants in the intervention group who had different levels of physical activity at baseline (≥7000 steps per day vs <7000 steps/day). The interaction between the physical activity level and intervention time point was analyzed in a linear mixed model adjusted for baseline age, sex, and frequency of app use. The data were analyzed using SPSS (version 29.0; IBM Corp). The statistical significance level was set at 5%.

**Results**

**Participants**

A total of 74 participants were initially enrolled in the study. However, 4 (5%) withdrew consent during the 12-week intervention period, with 1 (25%) withdrawal in the intervention group and 3 (75%) in the control group. The final analysis included 70 participants, 40 (57%) in the intervention group and 30 (43%) in the control group, in a modified intention-to-treat analysis, excluding those who withdrew consent (Figure 3). The intervention group comprised 9 groups of 4 to 5 participants each. A total of 5 (12%) out of 41 participants in the intervention group dropped out, including the 1 who withdrew consent.

---

https://aging.jmir.org/2024/1/e56184

JMIR Aging 2024 | vol. 7 | e56184 | p.23
"(page number not for citation purposes)"
**Participant Characteristics**

The baseline participant characteristics are shown in Tables 1 and 2. The mean age of the participants (n=70) was 77.3 (SD 6.1) years, with 26 (37%) male participants included in the study. There were no differences in baseline demographic characteristics between the intervention and control groups. However, participants in the intervention group were more likely to use apps and exercise more frequently. “Gather information and shop on the internet” for the intervention group, the control group and representative percentage were 78%, 57%, and 23.7%, respectively. “Use social networking services” for the intervention group, the control group and representative percentage were 60%, 37%, and 13.1%, respectively. “Do not use any information devices” for the intervention group, the control group and representative percentage were 0%, 3%, and 17%, respectively. Compared to the representative percentage of Japanese older adults based on the Annual Report on the Ageing Society published by the Japan Cabinet Office [32], both groups used ICT, with the intervention group exhibiting greater ICT use compared to the control group.

Although the difference in the baseline daily step count between groups was not statistically significant, the step count was higher in the intervention group than in the control group, with a median difference of >1000 steps. There was a statistically significant difference in baseline LPA and MVPA between the 2 groups; the intervention group exhibited significantly higher MVPA levels, while the control group showed significantly higher LPA levels. Furthermore, grip strength was higher in the intervention group, likely owing to the greater proportion of male participants; however, the difference was not statistically significant.

---

**Figure 3.** Flow diagram of participants. DPSA: digital peer support app.
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total sample</th>
<th>Intervention group (n=40)</th>
<th>Control group (n=30)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>77.3 (6.1)</td>
<td>76.9 (6.1)</td>
<td>77.9 (6.1)</td>
<td>.49a</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.12b</td>
</tr>
<tr>
<td>Male</td>
<td>26 (37)</td>
<td>18 (45)</td>
<td>8 (27)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>44 (63)</td>
<td>22 (55)</td>
<td>22 (73)</td>
<td></td>
</tr>
<tr>
<td>BMI (kg/m^2), mean (SD)</td>
<td>22.8 (2.9)</td>
<td>23.0 (2.9)</td>
<td>22.4 (2.9)</td>
<td>.43a</td>
</tr>
<tr>
<td>Living alone, n (%)</td>
<td>19 (27)</td>
<td>13 (32)</td>
<td>6 (20)</td>
<td>.22b</td>
</tr>
<tr>
<td>Self-rated health, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.13b</td>
</tr>
<tr>
<td>Excellent, good, or normal</td>
<td>63 (90)</td>
<td>38 (95)</td>
<td>25 (83)</td>
<td></td>
</tr>
<tr>
<td>Fair or poor</td>
<td>7 (10)</td>
<td>2 (5)</td>
<td>5 (17)</td>
<td></td>
</tr>
<tr>
<td>Perceived household economic status, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.57b</td>
</tr>
<tr>
<td>Excellent, good, or normal</td>
<td>67 (96)</td>
<td>39 (98)</td>
<td>28 (93)</td>
<td></td>
</tr>
<tr>
<td>Fair or poor</td>
<td>3 (4)</td>
<td>1 (2)</td>
<td>2 (7)</td>
<td></td>
</tr>
<tr>
<td>Life satisfaction, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.53b</td>
</tr>
<tr>
<td>Excellent, good or normal</td>
<td>62 (89)</td>
<td>35 (88)</td>
<td>27 (90)</td>
<td></td>
</tr>
<tr>
<td>Fair or poor</td>
<td>8 (11)</td>
<td>5 (12)</td>
<td>3 (10)</td>
<td></td>
</tr>
<tr>
<td>Working, n (%)</td>
<td>18 (26)</td>
<td>9 (22)</td>
<td>9 (30)</td>
<td>.48b</td>
</tr>
<tr>
<td>Smartphone ownership, n (%)</td>
<td>67 (96)</td>
<td>40 (100)</td>
<td>27 (90)</td>
<td>.07b</td>
</tr>
<tr>
<td>Frequency of app use, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.07b</td>
</tr>
<tr>
<td>Usually or sometimes</td>
<td>54 (77)</td>
<td>34 (85)</td>
<td>20 (67)</td>
<td></td>
</tr>
<tr>
<td>Rarely or never</td>
<td>16 (23)</td>
<td>6 (15)</td>
<td>10 (33)</td>
<td></td>
</tr>
<tr>
<td>Exercise habits, n (%)</td>
<td>37 (53)</td>
<td>24 (60)</td>
<td>13 (43)</td>
<td>.17b</td>
</tr>
<tr>
<td>Frequency of neighborhood interaction, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.53b</td>
</tr>
<tr>
<td>≥3 times/week</td>
<td>31 (44)</td>
<td>19 (48)</td>
<td>12 (40)</td>
<td></td>
</tr>
<tr>
<td>≤2 times/week</td>
<td>39 (56)</td>
<td>21 (52)</td>
<td>18 (60)</td>
<td></td>
</tr>
<tr>
<td>Participation in group exercise, n (%)</td>
<td>34 (49)</td>
<td>18 (45)</td>
<td>16 (53)</td>
<td>.49b</td>
</tr>
<tr>
<td>History falls in the past year, n (%)</td>
<td>9 (13)</td>
<td>5 (12)</td>
<td>4 (13)</td>
<td>.60b</td>
</tr>
<tr>
<td>Effect of the COVID-19 pandemic on decreased physical activity, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.64b</td>
</tr>
<tr>
<td>Great or slight</td>
<td>51 (73)</td>
<td>30 (75)</td>
<td>21 (70)</td>
<td></td>
</tr>
<tr>
<td>Not much or unchanged</td>
<td>19 (27)</td>
<td>10 (25)</td>
<td>9 (30)</td>
<td></td>
</tr>
<tr>
<td>Self-reported decrease in walking speed, n (%)</td>
<td>49 (70)</td>
<td>28 (70)</td>
<td>21 (70)</td>
<td>&gt;.99b</td>
</tr>
</tbody>
</table>

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

*b* Analysis was conducted using the chi-square test.
Table 2. Baseline measures of the triaxial accelerometer, physical function assessment, and self-efficacy for exercise (n=70).

<table>
<thead>
<tr>
<th>Outcome measures</th>
<th>Total sample</th>
<th>Intervention group (n=40)</th>
<th>Control group (n=30)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Triaxial accelerometer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steps/day, median (IQR)</td>
<td>5511 (3783-7852)</td>
<td>6310 (3936-8132)</td>
<td>5276 (3522-6275)</td>
<td>.08b</td>
</tr>
<tr>
<td>LPA (minutes/day), mean (SD)</td>
<td>329.9 (91.0)</td>
<td>306.6 (79.7)</td>
<td>360.3 (97.0)</td>
<td>.01d</td>
</tr>
<tr>
<td>MVPA (minutes/day), mean (SD)</td>
<td>45.5 (26.5)</td>
<td>51.0 (24.1)</td>
<td>38.4 (28.1)</td>
<td>.049d</td>
</tr>
<tr>
<td>SB (minutes/day), mean (SD)</td>
<td>538.2 (110.3)</td>
<td>530.1 (86.0)</td>
<td>522.3 (116.8)</td>
<td>.54d</td>
</tr>
<tr>
<td>Triaxial accelerometer wearing time (minutes/day), mean (SD)</td>
<td>913.6 (121.8)</td>
<td>888.4 (80.0)</td>
<td>946.4 (159.6)</td>
<td>.07d</td>
</tr>
<tr>
<td><strong>Physical function</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grip strength (kg), mean (SD)</td>
<td>25.6 (8.1)</td>
<td>26.9 (8.7)</td>
<td>23.8 (7.0)</td>
<td>.12d</td>
</tr>
<tr>
<td>CS-30, mean (SD)</td>
<td>20.7 (6.4)</td>
<td>20.7 (7.2)</td>
<td>20.8 (5.2)</td>
<td>.92d</td>
</tr>
<tr>
<td>Self-efficacy for exercise, mean (SD)</td>
<td>13.6 (3.6)</td>
<td>13.9 (3.6)</td>
<td>13.3 (3.6)</td>
<td>.54d</td>
</tr>
</tbody>
</table>

a Participants with valid accelerometer data; total sample, n=69; intervention, n=39; and control, n=30.
b Analysis was conducted using the Mann-Whitney U test.
c LPA: light intensity physical activity.
d Analysis was conducted using the independent samples t test (2-tailed).
e MVPA: moderate to vigorous intensity physical activity.
f SB: sedentary behavior.
g CS-30: 30-second chair-stand test.

Feasibility: Retention Rate, Number of Posts, Negative Impact, Continuation Factors, and Willingness to Continue

A total of 5 (12%) out of the 41 participants in the intervention group dropped out, resulting in a DPSA continuation rate of 88% (36/41). Reasons for dropping out included “withdrew research consent,” “not a good fit for me,” and “unknown cause” (n=1, 20% each), as well as “poor health” (n=2, 40%). The average number of total posts per person was 2.76 (SD 1.99) per day. After excluding participants who dropped out, the adherence rate was 96%, and the average number of total posts per person was 3.02 (SD 1.93) per day (Table 3).

Table 3. Digital peer support app adherence rate and average number of total posts per day among participants in the intervention group.

<table>
<thead>
<tr>
<th>Group</th>
<th>All participants (n=40)</th>
<th>Excluding participants who drop out (n=36)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (adherence rate, %)</td>
<td>Total posts/person/day, mean (SD)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n (adherence rate, %)</td>
</tr>
<tr>
<td>All</td>
<td>40 (87.7)</td>
<td>2.76 (1.99)</td>
</tr>
<tr>
<td>A</td>
<td>4 (95.8)</td>
<td>1.55 (0.32)</td>
</tr>
<tr>
<td>B</td>
<td>5 (88.3)</td>
<td>1.83 (0.88)</td>
</tr>
<tr>
<td>C</td>
<td>5 (80.4)</td>
<td>2.62 (1.72)</td>
</tr>
<tr>
<td>D</td>
<td>5 (79.8)</td>
<td>1.12 (0.59)</td>
</tr>
<tr>
<td>E</td>
<td>5 (77.1)</td>
<td>1.19 (1.72)</td>
</tr>
<tr>
<td>F</td>
<td>4 (99.7)</td>
<td>6.25 (1.46)</td>
</tr>
<tr>
<td>G</td>
<td>4 (98.6)</td>
<td>3.36 (1.16)</td>
</tr>
<tr>
<td>H</td>
<td>4 (95.5)</td>
<td>4.20 (1.07)</td>
</tr>
<tr>
<td>I</td>
<td>4 (74.1)</td>
<td>3.83 (2.88)</td>
</tr>
</tbody>
</table>

A total of 3 minor negative physical effects were reported; 2 participants reported knee pain and 1 reported plantar pain. There were no breaches of privacy associated with the use of the app. A total of 14 inquiries were received about how to use the app. The average response took approximately 15.4 (10.7) minutes per case. The reported reasons for continuing to use the app were continuing fellowship (28/36, 78%), having a common goal (24/36, 67%), having fun (20/36, 56%), tracking their step counts (20/36, 56%), maintaining motivation (18/36, 50%), sense of improved mental health (17/36, 47%), sense of improved physical health (14/36, 39%), being on the internet (14/36, 39%) and Emotional painless (11/36, 31%). Figure 4
presents the intention to continue using the DPSA after the 12-week intervention. The responses were strongly agree (11/36, 31%), somewhat agree (12/36, 33%), undecided (11/36, 31%), somewhat disagree (1/36, 3%), and strongly disagree (1/36, 3%). More than half of the participants (23/36, 64%) indicated an intention to continue.

**Figure 4.** Intention to continue using the digital peer support app after the end of the study among participants in the intervention group.

---

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

The results of the linear mixed model analysis of physical activity, physical function, and self-efficacy for exercise are listed in Table 4. There was a significant difference in the interaction between groups and intervention time points in the daily step count and MVPA (daily step count: $P=.04$ and MVPA: $P=.03$) but not in LPA, SB, grip strength, CS-30, and self-efficacy for exercise.

A post hoc subgroup analysis was conducted by dividing the intervention group into high-level physical activity and low-level physical activity subgroups [37] based on their level of physical activity at baseline. The step count, LPA, MVPA, and SB were compared (Table 5). The analysis showed a significant difference in the interaction between the groups and the daily step count at baseline ($P=.04$).
Table 4. Intervention effects on physical activity, physical function, and self-efficacy for exercise before and after the intervention.

<table>
<thead>
<tr>
<th>Outcome measuresa</th>
<th>Intervention group</th>
<th>Control group</th>
<th>Group/time (adjusted) P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Follow-up</td>
<td>Baseline</td>
</tr>
<tr>
<td>Accelerometer datab</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steps/day, median (IQR)</td>
<td>6310 (3936-8132)</td>
<td>8368 (5331-10,235)</td>
<td>5276 (3522-6275)</td>
</tr>
<tr>
<td>LPAc (minutes/day), mean (SD)</td>
<td>306.6 (79.7)</td>
<td>303.6 (91.0)</td>
<td>360.3 (97.0)</td>
</tr>
<tr>
<td>MVPA d (minutes/day), mean (SD)</td>
<td>51.0 (24.1)</td>
<td>65.7 (32.0)</td>
<td>38.4 (28.1)</td>
</tr>
<tr>
<td>SB (minutes/day), mean (SD)</td>
<td>530.1 (86.0)</td>
<td>522.3 (116.8)</td>
<td>522.3 (116.8)</td>
</tr>
<tr>
<td>Grip strengthf (kg), mean (SD)</td>
<td>26.9 (8.7)</td>
<td>26.0 (9.3)</td>
<td>23.8 (7.0)</td>
</tr>
<tr>
<td>CS-30g, mean (SD)</td>
<td>20.7 (7.2)</td>
<td>22.2 (7.3)</td>
<td>20.8 (5.3)</td>
</tr>
<tr>
<td>Self-efficacy for exerciseh, mean (SD)</td>
<td>13.9 (3.6)</td>
<td>14.8 (3.1)</td>
<td>13.3 (3.7)</td>
</tr>
</tbody>
</table>

aAnalyses were adjusted for age, sex, and frequency of app use (at baseline).
bParticipants with valid accelerometer data; intervention group: baseline, n=39, and follow-up, n=35; control group: baseline, n=30, and follow-up, n=30.
cLPA: light intensity physical activity.
dMVPA: moderate to vigorous intensity physical activity.
eSB: sedentary behavior.
fParticipants with valid grip strength data; intervention group: baseline, n=40, and follow-up, n=34; control group: baseline, n=30, and follow-up, n=23.
gCS-30; 30-second chair-stand test; participants with valid data; intervention group: baseline, n=40, and follow-up, n=32; control group: baseline, n=30, and follow-up, n=22.
hParticipants with valid self-efficacy for exercise data; intervention group: baseline, n=40, and follow-up, n=36; control group: baseline, n=30, and follow-up, n=28.

Table 5. Comparison of intervention effects on accelerometer data between low-level and high-level physical activity subgroups at baseline and follow-up in the intervention group.

<table>
<thead>
<tr>
<th>Outcome measuresa (accelerometer data)b</th>
<th>&lt;7000 steps/day (n=22)</th>
<th>≥7000 steps/day (n=18)</th>
<th>Group/time (adjusted) P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline</td>
<td>Follow-up</td>
<td>Baseline</td>
</tr>
<tr>
<td>Steps/day, median (IQR)</td>
<td>4338 (3207-5495)</td>
<td>5761 (4649-8680)</td>
<td>8581 (7571-10,117)</td>
</tr>
<tr>
<td>LPAc (minutes/day), mean (SD)</td>
<td>296.7 (78.1)</td>
<td>301.1 (94.8)</td>
<td>318.0 (82.3)</td>
</tr>
<tr>
<td>MVPA d (minutes/day), mean (SD)</td>
<td>33.1 (12.7)</td>
<td>52.0 (31.0)</td>
<td>71.9 (15.7)</td>
</tr>
<tr>
<td>SB (minutes/day), mean (SD)</td>
<td>538.2 (92.1)</td>
<td>513.3 (126.1)</td>
<td>522.4 (80.0)</td>
</tr>
</tbody>
</table>

aAnalyses were adjusted for age, sex, and frequency of app use (at baseline).
bParticipants with valid accelerometer data; low-level physical activity (<7000 steps/day) group: baseline, n=21, and follow-up, n=18; high-level physical activity (≥7000 steps/day) group: baseline, n=18, and follow-up (n=18).
cLPA: light intensity physical activity.
dMVPA: moderate to vigorous intensity physical activity.
eSB: sedentary behavior.

Discussion

Principal Findings

The study aimed to determine the characteristics of older adults who wanted to use the DPSA, which aimed to increase physical activity among older adults, and to confirm the feasibility of the DPSA and its impact on physical activity. Older adults who wanted to use the DPSA were more likely to be frequent users of the app and were more familiar with the use of ICT. Participants who reported an exercise habit tended to be more physically active at baseline. The retention rate was 88% (36/41) and the adherence rate was 87.7%, demonstrating the feasibility of older adults using the DPSA. The step count and MVPA increased significantly in the intervention group compared with those in the control group, demonstrating that the DPSA effectively increased physical activity. In DPSA users, participants with lower levels of baseline physical activity showed a more significant increase in their daily step count compared with those with higher levels of physical activity.
Comparison With Previous Studies

In this study, the retention and adherence rates were 88% (36/41) and 87.7%, respectively. These values are favorable compared to those reported in previous studies that have used digital technology to increase physical activity among older adults [38-42]. The findings from this study show that it is feasible for older adults to use DPSA to increase their level of physical activity. While differences were observed in adherence rates and the average number of total posts per day between groups, it is unclear what factors contribute to these differences. Only 3 negative physical effects were reported, but they were all minor and did not cause privacy breach issues. In contrast to our findings, Kullgren et al [43] showed that peer support using a 4-person web-based SMS text message board did not lead to an increase in physical activity among older adults. The authors attributed this lack of effectiveness to the failure to facilitate communication. In this study, the average number of comments per day in the intervention group was 2.76 overall and 3.01 excluding dropouts, indicating that many participants were actively using the DPSA. In addition, the fellowship was the factor with the highest percentage of intention to continue using DPSA. These may indicate that peer support based on social cognitive theory increased physical activity, as hypothesized. Self-efficacy is a key aspect of social cognitive theory [21]. However, in this study, although there was an increase in self-efficacy for exercise scores, the change was not significant. Possible reasons for the lack of a significant increase in self-efficacy for exercise in this study include the high baseline self-efficacy for exercise of the study population, the ceiling effect, and the short intervention period. In addition, the questionnaire used in this study may not reflect the impact of DPSA on self-efficacy. In this study, DPSA users showed a significant increase in their daily step count and MVPA duration despite the winter season. Participants with lower baseline physical activity levels showed a greater increase than those with higher levels of physical activity, suggesting that older adults with lower levels of physical activity may benefit more from using the DPSA than those with higher levels of physical activity. In the intervention group, the daily step count increased by >1000 steps on average. A systematic review of 17 prospective studies by Hall et al [44] has shown that each 1000-step increase in the daily step count decreases the risk of death and heart disease, with a 6% to 36% decrease in all-cause mortality risk and a 5% to 21% decrease in heart disease risk. Furthermore, it has been shown that an increase of 1000 steps per day decreases a woman’s risk of diabetes by 6% and an increase of 2000 steps per day decreases the risk of diabetes by 12% [45]. In this study, the duration of MVPA increased by >10 minutes. Previous studies conducted in the United States [46] have shown that adding 10 minutes per day of MVPA could prevent 6.9% of deaths per year in the US adult population aged between 45 and 85 years. A greater increase in physical activity is predicted to have a greater protective effect. In Japan, the Ministry of Health, Labour and Welfare published the ActiveGuide, the Japanese official physical activity guidelines for health promotion, in March 2013 [47]. The key message of this guideline is “+10,” indicating “add 10 minutes of MVPA per day.” We increased physical activity in older adults through a 5-year community-wide intervention that incorporated this guideline [48]. According to a meta-analysis of 26 cohort studies by Miyachi et al [49], an increase of 10 minutes of MVPA per day can cause a 3.2% reduction of the average relative risk of noncommunicable diseases, dementia, joint-musculoskeletal impairment, and mortality. The 2010 National Health and Nutrition Survey [50] found that 60.8% of the respondents are willing to take part in an additional 10 minutes of physical activity per day. Therefore, the “+10” recommendation could be feasible and efficient for the Japanese population [49]. On the basis of previous findings and the results of this study, DPSA is a viable and effective tool for increasing physical activity.

Limitations

This study has a few limitations. First, sampling problems such as small sample size and low statistical power, as well as the influence of confounding factors such as academic background, digital literacy, and motivation, cannot be ruled out. Indeed, there was a difference between the 2 groups regarding the percentage of information about the use of ICT. The enrolment to the intervention and control groups was nonrandomized, and participation was voluntary. A previous study by Tudor-Locke et al [37] has shown that healthy older adults tend to walk between 2000 and 9000 steps per day, with a median of 5500 steps. The control group had about the same number of steps as the average older adult in Japan, whereas the intervention group took approximately 1000 more steps per day than the average older adult in Japan. Both groups were highly interested in exercise, which might have influenced the effectiveness of the DPSA intervention. Second, this study used a short-term intervention period of 12 weeks. Previous studies have found that mHealth physical activity interventions are more effective over short periods of time (<16 and <12 weeks) than over longer periods of time and that the effects may not be maintained for longer periods of time [51,52]. Furthermore, other reviews have shown that mHealth interventions may promote small to moderate increases in physical activity and that the effects are maintained over time but that the effect size decreases over time [53]. In this study, 6% (2/36) of the participants stated that they did not have intention to continue the intervention after 12 weeks and 31% (11/36) answered undecided, indicating that high retention and adherence rates can be achieved for short-term use of 12 weeks but that the long-term retention rate is unknown. Third, the generalizability is limited, as participation in the intervention group was limited to those who owned a smartphone. In addition, they were already familiar with the use of ICT. In this study, training sessions were provided on the use of the app, so that even those who were less familiar with the use of the app could participate, but participation was low.

Conclusions

In this study, a 12-week intervention was conducted with older adults aged ≥65 years, using DPSA to increase physical activity. Older adults who used DPSA to increase physical activity tended to be more familiar with ICT and more physically active at baseline compared to the general older adult population. The feasibility of DPSA was good, with increases in daily steps and MVPA in the intervention group. Peer support–based interventions using digital apps may be effective in promoting...
physical activity among older adults. Notably, participants with lower levels of baseline physical activity showed a more significant increase in their daily step count compared with those with higher levels of physical activity. To confirm the effect of DPSA on physical activity and physical function in older adults, a randomized controlled trial should be conducted.

Acknowledgments
This study was supported by a Doctoral Student Grant-in-Aid Program from the Ushioda Memorial Fund. The authors would like to thank the staff of the Elderly Support Division of the Fujisawa City Government (Shinichi Sakae, Mayumi Taguchi, Mari Sakata, and Shizuka Kobayashi) for recruiting participants and managing the project; Ken Takahashi of the Fujisawa City Health and Medical Foundation for his cooperation as a health fitness instructor; Oguma laboratory staff for their help in collecting the data; and Professor Wataru Yamamoto for his guidance and support throughout the data analysis process. The authors would also like to thank Editage for English language editing.

Authors' Contributions
KT, YO, SY, and MS were responsible for designing this study. All authors were responsible for data collection. KT and YO were responsible for analyzing and interpreting the data. All authors edited, reviewed, and approved the final manuscript.

Conflicts of Interest
SY and MS were employees of A10 Lab Inc at the time of research.

Multimedia Appendix 1
The timing of the evaluation of each measurement item.

References

Multimedia Appendix 1
The timing of the evaluation of each measurement item.

References


Abbreviations

CS-30: 30-second chair-stand test
DPSA: digital peer support app
ICT: information and communication technology
LPA: light intensity physical activity
MET: metabolic equivalent
mHealth: mobile health
MVPA: moderate to vigorous intensity physical activity
SB: sedentary behavior
Designing Telemedicine for Older Adults With Multimorbidity: Content Analysis Study

Nida Buawangpong¹,², MD; Kanokporn Pinyopornpanish¹,²*, MD, PhD; Suphawita Pliaanuom¹,², MD; Nopakoon Nantsupawat¹,², MD; Nutchar Wiwatkunupakarn¹,², MD; Chaisiri Angkurawaranon¹,²*, MD, PhD; Wichuda Jiraporncharoen¹,²*, MSc, MD

¹Department of Family Medicine, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand
²Global Health and Chronic Conditions Research Group, Chiang Mai University, Chiang Mai, Thailand
*these authors contributed equally

Corresponding Author:
Wichuda Jiraporncharoen, MSc, MD
Department of Family Medicine
Faculty of Medicine
Chiang Mai University
Bunrueang Rit Rd. Suthep
Mueang Chiang Mai District
Chiang Mai, 50200
Thailand
Phone: 66 53935462
Email: wichuda.j@cmu.ac.th

Abstract

Background: Telemedicine is a potential option for caring for older adults with multimorbidity. There is a need to explore the perceptions about telemedicine among older adults with multimorbidity to tailor it to the needs of older adults with multiple chronic conditions.

Objective: This study aims to explore the perceptions about telemedicine among older patients with multimorbidity.

Methods: A qualitative study was conducted using semistructured interviews. The interview questions examined older adults’ perspectives about telemedicine, including their expectations regarding telemedicine services and the factors that affect its use. Thematic analysis was performed using NVivo (version 12; Lumivero). The study was reported using the Standards for Reporting Qualitative Research guidelines.

Results: In total, 29 patients with multimorbidity—21 (72%) female patients and 8 (28%) male patients with a mean age of 69 (SD 10.39) years—were included. Overall, 4 themes and 7 subthemes emerged: theme 1—perceived benefit of telemedicine among older adults with multimorbidities, theme 2—appropriate use of telemedicine for multimorbid care, theme 3—telemedicine system catering to the needs of older patients, and theme 4—respect patients’ decision to decline to use telemedicine.

Conclusions: Telemedicine for older adults with multimorbidity should focus on those with stable conditions. This can help increase access to care for those requiring continuous condition monitoring. A structured telemedicine program and patient-centered services can help increase patient acceptance of telemedicine. However, health care providers must accept the limitations of older patients that may prevent them from receiving telemedicine services.

(JMIR Aging 2024;7:e52031) doi:10.2196/52031

KEYWORDS

telemedicine; telehealth; chronic disease; multimorbidity; older adults; mobile phone
Introduction

Background

The COVID-19 pandemic has led to the emergence of telemedicine as a viable alternative to traditional, in-person care. Telemedicine has the potential to provide convenient medical care for patients with disabilities, transportation limitations, or busy schedules, enabling them to receive care from home [1-3]. It has become an increasingly valuable tool for delivering care to patients with multimorbidity, who require regular monitoring or adjustments to their treatment plans [4,5]. An emerging care model illustrates the integration of a patient-centered approach for individuals and chronic care model with multimorbidity. The model offers comprehensive care across various patient aspects and uses a multidisciplinary approach to address the complexity of managing multimorbidity [6]. Telemedicine can provide remote consultations and monitoring, provide patient education, and facilitate continuity of care [7,8].

In many countries, telemedicine has been promoted in the post–COVID-19 era owing to comparable health outcomes and favorable cost-effectiveness compared with in-person visits [9]. In Thailand, telehealth projects have been launched by the National Broadcasting and Telecommunication Commission of Thailand and the Thai Ministry of Public Health to improve health care services’ accessibility and quality [10]. As Thailand has become an aged society, with approximately 12 million people aged ≥60 years [11], telemedicine can help improve the quality of life of older adults by promoting healthy behavior, enhancing social functioning, and reducing depressive symptoms [12,13]. Telemedicine can also improve health for older patients, who often have multimorbidity, by helping to provide continuous medical care [14].

However, the lessons learned from using telemedicine during the COVID-19 pandemic have documented challenges for both older patients and health care providers (HCPs), as it was an unfamiliar mode of treatment compared with in-person service [15]. In addition, the literature suggests that older patients may prefer in-person visits owing to the frustration caused by technological challenges when using telemedicine [16]. However, their perceptions and preferences regarding telemedicine remain poorly understood. Therefore, the needs of older patients for telemedicine should be explored [17].

Moreover, there are gaps in understanding telemedicine for those with multimorbidity [6,18]. Many studies showed the effectiveness of telemedicine, but most focused on the use of telemedicine for a single disease [19-21]. Managing patients with multimorbidity is an increasing challenge in primary care practice [22]. In multimorbidity, there are many interactions such as disease-disease, treatment-treatment, and disease-treatment, which increase the complexity of management [23,24]. The rising question is what is the appropriate use of telemedicine in caring for older patients with multimorbidity [25].

Objective

Our study aimed to explore the perceptions about telemedicine among older patients with multimorbidity. By understanding the perceptions about expectations, preferences, and barriers regarding telemedicine, the results can be used to develop telemedicine strategies to support the management of multimorbidity in the older population.

Methods

Study Design

A qualitative study of older adults with multimorbidity attending a primary care outpatient clinic at a university hospital in Thailand was conducted in 2021. The study used the Standards for Reporting Qualitative Research guidelines, which is a list of 21 items considered essential for complete, transparent reporting of qualitative research [26].

Setting and Participants

The Family Medicine Clinic at the Faculty of Medicine, Chiang Mai University, is a primary care clinic that provides general medical care for chronic diseases. Approximately 80% of all patients are older adults (aged ≥60 y) with multimorbidity. Multimorbidity was defined as the patient’s illness that includes the presence of multiple diseases or conditions, often with a cutoff of ≥2 conditions [27]. The most common conditions are hypertension, type 2 diabetes mellitus, and dyslipidemia. Approximately 90% of the patients had multimorbidity. Among those attending the clinic, the rate of controlled hypertension, type 2 diabetes mellitus, and dyslipidemia between August and October 2021 were 93.7%, 70.3%, and 79.8%, respectively.

Convenience sampling was used. Patients aged ≥60 years and diagnosed with at least 2 chronic conditions who were accessible and available were invited to participate in the study. The patients needed to have stable conditions, defined as being asymptomatic and not having any urgency or emergency conditions, according to the national clinical guidelines for managing hypertension [28], type 2 diabetes mellitus [29], and dyslipidemia [30].

Telemedicine Service

Telemedicine is the use of electronic information and communication technologies to provide and support health care when distance separates the participants [31]. In 2020, the Family Medicine Clinic in Thailand started a telemedicine service in response to the COVID-19 pandemic. After the COVID-19 pandemic, our facility continues to use telemedicine for delivering care to patients with chronic conditions. By considering patient safety, patients with moderate to well-controlled chronic conditions were approached to participate in the telemedicine service for continuous care [32]. In addition, it is essential to discuss with the patients the purpose of telemedicine and to address any limitations associated with its use [33]. In addition, the objective of telemedicine as ongoing care and limitations of telemedicine need to be discussed with the patients.

The clinic had adopted a published multimorbidity assessment checklist developed to help care for patients with multimorbidity.
The 20-item assessment checklist (Simple Multimorbidity Assessment Checklist for Primary Care) incorporates patient-centered concepts into managing multimorbidity in primary care settings, including assessment of the patient, review of all diseases and conditions, review of all treatments, review of clinical practice guidelines, assessment of interactions, understanding patient context and concerns, finding common ground, setting individual care plan, and continuity of care and follow-up visits. The checklist was also extended for use in the telemedicine service for assessing patients who were suitable for telemedicine and health caregiving.

The telemedicine service was provided via video or audio call, depending on the patients’ available devices and abilities. Physicians could collect patients’ medical histories and evaluate emergency or urgent conditions during the consultation. If any patient had conditions that required further evaluation, they were advised to come to the hospital. If patients did not require any further in-person assessment, medications were prescribed and delivered to the patient’s home via post.

Ethical Considerations
This study was reviewed and approved by the institutional ethics committee of the Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand (approval number 227/2021). All participants were informed about the research study and provided consent.

Data Collection and Analysis
Semistructured interviews were conducted between September and November 2021. The interview questions (Multimedia Appendix 1) focused on older patients’ perceptions about telemedicine, including their preferences regarding expectations from telemedicine services, and factors that affect its use, using the Unified Theory of Acceptance and Use of Technology framework [34]. This framework illustrates a comprehensive understanding of all the factors that affect people’s intentions to use the new technology. A research assistant, not involved in providing medical care, was trained in the interview method and interview questions by WJ and NB. Each interview lasted approximately 20 minutes and was conducted on-site. Interviewed information included baseline characteristics (age, sex, educational level, employment status, and their decision regarding telemedicine services) with permission for audio recording. Then, the interviews were transcribed verbatim. Data collection and analysis were performed iteratively by researchers. Recruitment ended when data saturation of the core analytic content had been achieved. Previous literature suggested that the sample size of 9 to 17 interviewees could help to reach saturation [35]. We further determined the sample size based on a previous study investigating the crucial factors for outpatient service selection among older adults. At least 16 patients were required to achieve data saturation [36]. Therefore, we considered collecting data from at least 16 patients until we achieved data saturation in the results [37].

Each transcript was evaluated multiple times to aid familiarization and understanding of the data. Descriptive analysis was used to describe the patient’s characteristics. For qualitative analysis, 2 independent researchers (NB and WJ) conducted inductive thematic analysis [38]. The preliminary results were then interpreted and discussed with KP, SP, NN, and CA. Codes were then developed based on patterns in the data. The identified codes were compared and discussed for similarities and differences until consensus was reached regarding the emergent themes and subthemes. Data analysis was performed using NVivo (version 12; Lumivero).

Results
Overview
In total, 29 older patients with multimorbidity participated. Of the 29 participants, 21 (72%) were women and 8 (28%) were men. The mean age was 71 (SD 7.17) years. The 2 most prevalent underlying diseases were dyslipidemia (27/29, 93%) and hypertension (25/29, 86%). Most patients had completed primary school (11/29, 38%) or had a bachelor’s degree (10/29, 35%). Of the 29 patients, 23 (79%) were retired and 6 (21%) were self-employed. Of the 29 patients, 18 (62%) patients were interested in using telemedicine, whereas 11 (38%) patients were not interested in telemedicine and rejected telemedicine when they were offered. Patients’ characteristics are summarized in Table 1.

From the semistructured interviews, 4 themes and 7 subthemes emerged. The themes and subthemes are summarized in Textbox 1.
Table 1. Patients' characteristics (N=29).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (y), mean (SD)</strong></td>
<td>71 (7.17)</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>21 (72)</td>
</tr>
<tr>
<td>Male</td>
<td>8 (28)</td>
</tr>
<tr>
<td><strong>Chronic conditions, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Dyslipidemia</td>
<td>27 (93)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>25 (86)</td>
</tr>
<tr>
<td>Type 2 diabetes mellitus</td>
<td>9 (31)</td>
</tr>
<tr>
<td>Others</td>
<td>6 (21)</td>
</tr>
<tr>
<td><strong>Number of chronic conditions, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>19 (66)</td>
</tr>
<tr>
<td>3</td>
<td>7 (24)</td>
</tr>
<tr>
<td>4</td>
<td>1 (3)</td>
</tr>
<tr>
<td>5</td>
<td>2 (7)</td>
</tr>
<tr>
<td><strong>Educational level, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>No education</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Primary school</td>
<td>11 (38)</td>
</tr>
<tr>
<td>Secondary school</td>
<td>5 (17)</td>
</tr>
<tr>
<td>Vocational certificate</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>10 (34)</td>
</tr>
<tr>
<td>Master’s degree</td>
<td>1 (3)</td>
</tr>
<tr>
<td><strong>Working status, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Retired</td>
<td>23 (79)</td>
</tr>
<tr>
<td>Self-employed</td>
<td>6 (21)</td>
</tr>
<tr>
<td><strong>Decision regarding telemedicine, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Accept</td>
<td>18 (62)</td>
</tr>
<tr>
<td>Decline</td>
<td>11 (38)</td>
</tr>
</tbody>
</table>

Textbox 1. Summary of themes and subthemes.

**Theme 1: perceived benefit of telemedicine among older adults with multimorbidities**
- Convenient to access without the need for travel
- Minimize the risk of COVID-19 transmission

**Theme 2: appropriate use of telemedicine for multimorbidity**
- Telemedicine for monitoring stable conditions
- Enhancing the self-management of chronic conditions

**Theme 3: telemedicine system catering to the needs of older patients**
- Telemedicine services should be as similar as possible to in-person care
- Telemedicine services should adopt a clear protocol that includes in-person visits
- Supporting the development of technological skills and providing resources

**Theme 4: respecting patients’ decision to decline to use telemedicine for various reasons**
Theme 1: Perceived Benefit of Telemedicine Among Older Patients With Multimorbidity

Overview

The benefits of telemedicine were collected from older patients. They reported the reasons why telemedicine should be used and its benefits. These include eliminating the requirement for travel and reducing the risk of contracting COVID-19.

Subtheme 1: Convenient to Access Without the Need for Travel

The most mentioned benefit of telemedicine was eliminating the need for travel. Participants mentioned that telemedicine is convenient for accessing and receiving continuity of care. It can also save time and money, such as time spent in driving and finding parking. A patient mentioned that telemedicine reduces stress from long wait times at hospitals. It also eliminates the risk of driving accidents, especially in older patients with sensory problems owing to physiologic changes. In addition, there is no burden on family members or caregivers to come and drop them at the hospital:

...If we compare the advantages and disadvantages, there are more advantages, as it saves both time and cost. We don’t have to drive, look for parking spots, or wait in line to meet the doctor and receive medication. The advantages are greater. [Participant 13; female; aged 64 years; teacher; uncontrolled hypertension and obesity]

I like it because I don’t have to go to the hospital. It’s convenient. If I had to give [telemedicine] a score, it would be a ten because it’s convenient for me. I don’t have to drive there because I’m not good at driving right now. I have to ask my husband to take me there. [Participant 5; female; aged 72 years; housemate; well-controlled hypertension, type 2 diabetes mellitus, and dyslipidemia]

Because we don’t have to go to the hospital anymore. They send the medication to our house. It’s difficult to go to the hospital now. We have to ask our children to take us, but everyone is working. I want to receive the medication at home because I take this medication regularly. [Participant 26; female; aged 72 years; retired; well-controlled hypertension, type 2 diabetes mellitus, and dyslipidemia]

Sometimes, if I go to the hospital to see a doctor, I have to wait a long time, and it can be stressful. With video calls, I can talk for a long time. [Participant 23; male; aged 76 years; retired; uncontrolled type 2 diabetes mellitus, well-controlled hypertension, and dyslipidemia]

Subtheme 2: Minimize the Risk of COVID-19 Transmission

Patients perceive telemedicine as a helpful way to reduce the risk of SARS-CoV-2 infection by avoiding contact with individuals with infection at the hospital. By not having to physically go to the hospital, there is no need to wait in crowded areas for a physician or medication after treatment, resulting in decreased rate of contact:

I don’t have to go to places with many COVID-19 cases. My daughter also likes it because I don’t have to take risks. Using telemedicine is very good for me. [Participant 2; female; aged 72 years; self-employed; well-controlled hypertension, type 2 diabetes mellitus, and dyslipidemia]

Theme 2: Appropriate Use of Telemedicine for Multimorbidity

The participants felt that telemedicine should be used to care for patients with stable conditions in evaluating, monitoring, and providing health promotion. Participants also acknowledged the limitations of telemedicine in providing medical care, such as the inability to perform a complete physical examination or blood tests.

Subtheme 1: Telemedicine for Monitoring Stable Conditions

Patients feel confident in receiving telemedicine services when they have stable conditions because they have no abnormal symptoms, and the on-site care provided is only in the form of conversation to monitor their condition. Telemedicine services for those with stable conditions can resemble on-site care. Telemedicine services should include monitoring of clinical symptoms, vital signs, body weight, and behavioral factors. Consultation time is also required, so that patients can consult with their physician and inquire about their condition and receive follow-up care:

It’s just like when we see a doctor at the hospital. If we meet the doctor, we ask questions like this. We can also ask online like this and see each other’s faces; finding a doctor this way is good and convenient. The doctor called, and we talked. If we have any questions, we ask, and the doctor answers. It’s just like going to see a doctor. [Participant 18; female; aged 69 years; retired; well-controlled type 2 diabetes mellitus and osteoporosis]

However, some patients still believe that if they experience new or more severe symptoms, they prefer to receive treatment in an in-person setting for more detailed examinations or blood tests:

It’s [telemedicine] comprehensive, but only if I do not have severe symptoms... However, if the patient has more severe symptoms..., it’s uncertain how effective the treatment [received through telemedicine] will be. [Participant 19; female; aged 66 years; retired; well-controlled hypertension and type 2 diabetes mellitus]

It’s [telemedicine] good. Luckily, there have been no issues during this period. But if any problems arise, I still have to go see a doctor. [Participant 15; female; aged 73 years; retired; uncontrolled type 2 diabetes mellitus and well-controlled hypertension]
Subtheme 2: Enhancing the Self-Management of Chronic Conditions

The patients perceive that they can take better care of themselves when telemedicine provides health care information specific to their health problems. Some patients suggested incorporating self-monitoring and health promotion features into the telemedicine platform. They believe in sharing self-management information with physicians to improve disease management, such as home monitoring of blood pressure and blood sugar levels. The platform could also provide specific knowledge for lifestyle modification, such as exercise videos or electronic brochures about food exchange lists. This enhances the potential for self-management:

> I normally check and record my blood sugar and blood pressure at home. Sometimes, I forget to bring the records to the hospital. However, in telemedicine, when the doctor calls, I can inform them of my records. [Participant 2; female; aged 72 years; self-employed; well-controlled hypertension, type 2 diabetes mellitus, and dyslipidemia]

> I want to know more about blood pressure. We may think it’s not a big deal, but actually, it’s a silent danger. The doctor said it’s a scary disease, and I want information on how to take care of this disease. It would be great if there were some tips on this matter. [Participant 14; female; aged 66 years; government employee; uncontrolled hypertension and well-controlled type 2 diabetes mellitus]

> If there is a LINE group [chat group], I would like to receive informative messages about health. Even seniors need to read news and information related to health, such as knowledge about [COVID] vaccines. [Participant 21; female; aged 75 years; retired; well-controlled hypertension and type 2 diabetes mellitus]

Theme 3: Telemedicine Services Catering to the Needs of Older Patients

For an effective telemedicine service for older patients, it is crucial to address patient concerns and establish a clear management protocol based on their health status. Providing supporting resources, having good communication skills, and being aware of potential barriers arising from unfamiliarity with technology are also essential in meeting their needs.

Subtheme 1: Telemedicine Services Should Be as Similar as Possible to In-Person Care

When using telemedicine, patient concerns must still be evaluated, similar to in-person service. Some concerns may persist even after receiving telemedicine services. Presenting conditions can influence their physical or mental well-being, making it crucial for HCPs to thoroughly understand patients’ illnesses to ensure appropriate management. Some patients feel that video call feels more similar to an actual on-site visit than audio calls because they can see the facial expressions and gestures of the HCPs:

> Video calls would be better because the doctor can see the patient’s face and how they feel at that moment, whether they are feeling stressed or not. With video calls, I feel closer [to the doctor] and more comfortable.... I usually don’t share things within my family unless it’s with the doctor. But with video calls, I feel more comfortable because I can see the doctor’s face, knowing that they care about me. I just want the doctor to call me and ask what I want to share or talk about. [Participant 1; female; aged 62 years; retired; well-controlled hypertension and type 2 diabetes mellitus]

> With video calls, we feel close to each other. It feels like we’re still talking to each other. It’s good because we can talk to the doctor about anything comfortably. I think it’s a good thing because patients can express themselves fully to the doctor. [Participant 17; male; aged 76 years; retired; well-controlled hypertension and type 2 diabetes mellitus]

Subtheme 2: Telemedicine Services Should Adopt a Clear Protocol That Includes In-Person Visits

Patients express concerns that telemedicine might replace traditional, in-person care, leading to a lack of access to physical examinations, blood tests, and additional symptom management. They desire telemedicine to complement a comprehensive multimorbidity management program while still having the option to see physicians in person at the hospital:

> Sometimes I want to meet with the doctor in-person to talk directly or ask questions. The doctor can know my symptoms if I communicate directly. If I say that it hurts here or it is swollen here, the doctor can touch it and examine it for evaluation. This is the basic step of diagnosing symptoms. I mean, I want to meet the doctor sometimes, but not frequently. It’s not like I don’t see the doctor for a year. I just want to see the doctor once or twice to feel reassured. [Participant 16; female; aged 74 years; retired; well-controlled hypertension, type 2 diabetes mellitus, and dyslipidemia]

The support system of telemedicine services was crucial for patients’ decision-making regarding whether to accept or decline the service. Most patients are willing to accept telemedicine owing to clear operational systems, including appointment scheduling, notifications, web-based payment, and medication delivery. However, some patients still have doubts about payment systems and medication delivery. In addition, patients receiving telephone-based care may have uncertainties about the authenticity of the HCP:

> After a case manager added me on LINE [application], they gave me an appointment for a video call. A day before the appointment, a nurse called me and said the doctor would have a video call tomorrow. When the appointment arrived, they would call me and ask if it was convenient for the doctor to have a video call now. I answered that it was. Then, the doctor called me. It was a very good process. [Participant 15; female; aged 73 years; retired;
uncontrolled type 2 diabetes mellitus and well-controlled hypertension]

It’s possible to send the appointment time through LINE in advance; for example, if the doctor would come in the afternoon, someone would call in the morning to inform. The doctor could then ask about the symptoms very well. It was done quickly, in just a moment. [Participant 21; female; aged 75 years; retired; well-controlled hypertension and type 2 diabetes mellitus]

I asked about the cost of the medicine that sent text information by phone because I couldn’t contact the finance department. It’s very difficult. [Participant 18; female; aged 69 years; retired; well-controlled type 2 diabetes mellitus and osteoporosis]

I don’t know if they are doctors or not. To be honest, I don’t know who is calling me. But if it’s a video call, I can be more confident. [Participant 13; female; aged 64 years; teacher; uncontrolled hypertension and obesity]

Subtheme 3: Supporting the Development of Technological Skills and Providing Resources

Most participants had limited technological skills and relied on their children to assist in using electronic telemedicine devices. Only a small minority were proficient in using such devices, with some preferring mobile phone calls over video calls owing to incompatible smartphones or unfamiliarity with more complex devices such as tablets or PCs. However, some older individuals expressed willingness to learn with proper support:

If necessary, I need to adapt. I have to learn to use additional equipment because I don’t usually have a smartphone, so it might be difficult to learn. I can’t even turn it on. [Participant 9; female; aged 78 years; retired; well-controlled hypertension, type 2 diabetes mellitus, and dyslipidemia]

If it’s time for telemedicine, I have to try to adapt and learn gradually. [Participant 10; female; aged 71 years; retired; well-controlled hypertension, type 2 diabetes mellitus, and dyslipidemia]

Theme 4: Respecting Patients’ Decision to Decline to Use Telemedicine for Various Reasons

Older patients often reject telemedicine owing to various obstacles. They face challenges related to age-related physiological changes, including forgetfulness and cognitive difficulties such as finding phones or using video calls despite instructions. In addition, some patients do not regularly use electronic devices, whereas others feel burdensome relying on their children for telemedicine assistance. A few patients are unable to receive telemedicine service owing to their routine work commitments. Therefore, HCPs need to understand these obstacles and respect their decisions for declining telemedicine:

I don’t want to use telemedicine. It’s not difficult for me to see a doctor in person. If you teach me something, I will forget in three months. For example, when I wanted to take a video, my grandchild had to teach me ten times, but when I got home, I couldn’t remember. My memory has not been good for a few years. [Participant 12; male; aged 68 years; self-employed; well-controlled hypertension and type 2 diabetes mellitus]

Sometimes I am not with my phone. Like when I went to a restaurant, I forgot my phone there and didn’t realize it for three days. I think it’s a problem related to age, but if I were newly retired at 60-65 years old, I would be fine. But now that I’m nearly 80, I have problems, especially with memory and internet use. [Participant 23; male; aged 76 years; retired; well-controlled hypertension and type 2 diabetes mellitus]

If my child can help, that would be great. I need my child to be here because I don’t know much. I’m forgetful, but I don’t know if my child is available to help or not. I can’t do it if I’m alone because I have to care for two other older people who are 90 years old. It’s not easy for me because I have to take care of others as well. [Participant 4; female; aged 63 years; retired; uncontrolled type 2 diabetes mellitus, well-controlled hypertension, and dyslipidemia]

I can participate, but I’m not familiar with it. I’m old and have never used LINE [chat application] before. Trying to learn it now may be difficult because my memory is not very good, and I tend to forget things easily. Although my grandchildren have computers and mobile phones, I don’t want to bother them because they have to work all the time. If the doctor needs to call me, I have to ask my children. I don’t know if they’re available to answer or not. [Participant 8; female; aged 71 years; retired; uncontrolled type 2 diabetes mellitus and well-controlled dyslipidemia]

Discussion

Principal Findings

Summary

In this qualitative study, participants perceived telemedicine as beneficial because it eliminates the need for travel and minimizes the risk of COVID-19 transmission. Older adults view telemedicine as a safe and effective way to manage stable chronic conditions. It is recommended to include health promotion in telemedicine services to enhance self-management. Regarding catering to older patients’ needs, patients expressed that telemedicine should be presented as part of a continuous care program for multimorbidity, incorporating web-based monitoring with periodic in-person visits for physical examinations and laboratory screenings. Clear instructions, technological skills training, and access to resources such as equipment and caregivers are essential to make the program user-friendly for older patients. However, HCP should respect patients’ decision to decline telemedicine owing to various obstacles that older patients may face when using it.
Patients accepted telemedicine as an effective method to improve access to health care for older patients. Some older patients face difficulties when coming to hospitals, such as finding transportation and parking space, long waits for the physician, and long queues for receiving medication [16,36]. This is in accordance with the literature that positive perceptions about telemedicine include cost savings [39]. During the COVID-19 pandemic, receiving treatment through telemedicine services also helped to reduce the risk of infection transmission by reducing the risk of overcrowding of patients and the risk of exposure for those who may not need to come to the hospital [40]. Telemedicine would be a necessary solution for addressing problems regarding access to care in the event of new pandemics.

Several studies have shown that telemedicine can effectively improve the health care outcomes of older patients, particularly those with chronic conditions such as diabetes [19], heart disease [20], and asthma [21]. A study found that telemedicine consultations reduced hospitalization rates among older patients with chronic heart failure compared with standard care [41]. Another study found that telemedicine consultations for older patients with chronic obstructive pulmonary disease improved symptom control and quality of life and reduced hospitalization rates [42]. There is evidence supporting that telemedicine services have the potential to enhance self-management among patients and their families, including improving medication adherence among older patients with chronic conditions and improving disease control and patient satisfaction [43,44].

Older patients, who often have >1 chronic condition, constitute a key group who use the health systems [45]. Owing to the situation in Thailand, it will be a superaged society in the next few decades [46]. The older adults would be the main target for health care delivery. There are potential opportunities in digital health such as telemedicine, emphasizing the management of chronic diseases in Thailand [47]. Cost-effectiveness was also another reason in the long run for telemedicine compared with an in-person visit [4,48]. On the basis of patients’ perceptions obtained from our study and previous evidence supporting the health outcomes of telemedicine, we have the following 4 suggestions for enhancing telemedicine services tailored to the needs of older patients with multimorbidity.

**Identify the Target Population as Individuals With Stable Chronic Conditions**

Our study found that telemedicine is an accepted model for promoting continuous care for older patients with multimorbidity. In cases where patients have well-controlled chronic conditions and no abnormal symptoms, they can receive symptom monitoring and treatment through telemedicine [49]. Health care services for older patients with stable chronic diseases may not need to differ between telemedicine and in-person visits. The services should aim to monitor patients with stable conditions by regularly inquiring about their symptoms; offering self-care instructions at home (such as measuring blood pressure and blood sugar levels); and encouraging healthy behaviors such as medication adherence, maintaining a proper diet, and engaging in regular exercise. Telemedicine has the potential to replace nonurgent in-person medical visits for stable chronic diseases, as it can be used for symptom monitoring, detecting complications or disease progression, and prescribing medication delivery for stable chronic conditions [7,50].

**Telemedicine Services Should Be Designed to Closely Resemble In-Person Visits With Scheduled Periodic In-Person Visits**

Patients expressed concerns regarding the quality of care and maintaining the physician-patient relationship received through telemedicine. A previous study revealed that older patients perceived in-person visits as fostering a strong physician-patient relationship compared with telemedicine [16]. To address these concerns, telemedicine services should strive to deliver care that closely resembles an in-person visit in terms of the process and pattern of care. Telemedicine using video calls closely simulates an in-person visit over phone calls. It helps reduce medication errors, enhances diagnostic accuracy, and improves decision-making accuracy [51]. Telemedicine holds the potential to facilitate shared decision-making between patients and HCPs, thus promoting a patient-centered approach to care [52]. HCPs can also leverage telemedicine to provide education, promote behavior change, empower patients to take control of their health, boost their confidence, and ensure continuity of care [53].

Nevertheless, it is important to acknowledge the limitations of telemedicine. It is unable to perform comprehensive physical examinations and detailed laboratory tests, which means that it cannot fully replace in-person visits. Hence, it is crucial to integrate regular in-person visits with a physician at a hospital to adhere to standard medical practices. These in-person visits can be scheduled periodically on an annual basis, which can help instill confidence and satisfaction with the telemedicine services [54].

**Integrated Support Systems for Telemedicine, Including Clear Protocols, Caregiver Assistance, and Electronic Health Literacy Training**

When caring for older patients with multimorbidity through telemedicine, it is essential to establish a program that adheres to standard practices, incorporates clear protocols, and provides the necessary technical skills and resources. This will help ensure that telemedicine remains as a viable option beyond the COVID-19 pandemic, emphasizing ease of use and demonstrating its benefits. A well-defined service program enables patients to understand the process of care they will receive and empowers them to communicate their specific health needs [55]. In addition, telemedicine services for older adults may require assistance at various stages, including guidance for using tools such as smartphones, instruction for use, and involving caregivers in the process [56]. These supports could contribute to a smooth and more effective telemedicine experience for older patients and their caregivers.

**Develop Alternative Services for Older Adults Who Cannot Use Telemedicine**

Despite the potential benefits of telemedicine for the care of older adults, its use has some challenges and limitations. Some
older patients may face various obstacles in using telemedicine, such as declining vision, hearing, and memory owing to aging; difficulty in learning new skills; unfamiliarity with technology; and feeling burdened to ask for support, which can lead to rejecting telemedicine services [56]. In addition, telemedicine is unsuitable for unstable patients requiring emergent management and detailed physical examinations that cannot be conducted remotely [57]. Some specific clinical contexts or onset of new symptom in multimorbidity, such as hemiparesis, require a comprehensive examination for critical diagnostic accuracy and severity evaluation. If possible, an in-person visit would be more appropriate [58]. Thus, in-person visits or other alternatives should be available [7].

**Strengths and Limitations**

The study has several strengths and limitations. The strength of this study lies in its structured approach to gathering insights about various aspects of telemedicine service tailored to the need of older adults with multiple chronic conditions. The study results can provide valuable guidance about preparing and delivering telemedicine services for this population. However, there are still some limitations to be considered. First, participants were recruited from a single health care facility. The results may be affected by the nature of the health care system and the educational level and digital literacy level of the population. Further studies from different settings and regions are needed to tailor telemedicine services to the needs of older adults with multimorbidity. In addition, future studies could explore more experiences of HCPs providing telemedicine services to this population and identify strategies to address their challenges and concerns. Next, we did not include uncontrolled conditions in this study, and we did not include health care professionals. Further studies may include other telemedicine users. Another consideration point that could influence the results of this study is gender. Gender is associated with differences in digital health care behavior and plays a role in the adoption of health technologies. For example, women exhibit high tendency to access health care services, book physician’s appointments, and search for nutrition-related information. In contrast, men are more likely to explore options related to smoking cessation and use health apps for monitoring sleep patterns and blood pressure than women [59].

As there are various guidelines for telemedicine management, the platform used depends on the facility in each hospital setting. Having many telemedicine providers may disrupt the continuity of care owing to regulation and personal data protection [60]. Furthermore, there remains an inequity in access to care in telemedicine for vulnerable population, such as older people. A strategy to promote electronic health education and provide the necessary equipment to ensure telemedicine equity is needed [61]. Therefore, the use of telemedicine should be tailored depending on the setting and needs of the population and health system.

To ensure the sustainability of telemedicine for older patients with multimorbidity, HCPs should consider patients’ needs, expectations, and abilities when designing telemedicine systems. Importantly, the findings also suggest that HCPs should respect the decision of older patients who decline to use telemedicine owing to multiple obstacles and find alternative ways.

**Conclusions**

The study highlights the importance of personalized and patient-centered care [62], where providers should understand older adults’ needs, preferences, and limitations to tailor telemedicine services for the population. The use of telemedicine for older patients with multimorbidity should focus on those with stable conditions. For this population, telemedicine can help increase access to medical services for patients who require continuous monitoring and care. A structured program incorporating periodic in-hospital visits can help increase patient acceptance of telemedicine. However, HCPs must also understand the limitations of older patients owing to various factors that may prevent them from receiving telemedicine services.

**Acknowledgments**

This study was supported by a grant from the Faculty of Medicine, Chiang Mai University (grant number 043/2563), and partially funded by Chiang Mai University. The funders had no role in the study’s design; collection, analyses, or interpretation of data; writing of the manuscript; or decision to publish the results.

**Conflicts of Interest**

None declared.

Multimedia Appendix 1

Interview questions.

[DOCX File, 13 KB - aging_v7i1e52031_app1.docx]

**References**


Abbreviations

HCP: health care provider
Examining the Effect of Contactless Intergenerational Befriending Intervention on Social Isolation Among Older Adults and Students’ Attitude Toward Companionship: Content Analysis

Keya Sen¹, PhD; Nida Laheji¹, MHA; Zo Ramamonjiarivelo¹, PhD; Cecil Renick¹, JD; Randall Osborne², PhD; Brad Beauvais¹, PhD

¹School of Health Administration, Texas State University, San Marcos, TX, United States
²Department of Psychology, Texas State University, San Marcos, TX, United States

Corresponding Author:
Keya Sen, PhD
School of Health Administration
Texas State University
601 University Drive
San Marcos, TX, 78666
United States
Phone: 1 512 245 3508
Email: Keyasen@txstate.edu

Abstract

Background: Intergenerational friendship, a mechanism of social support, is an effective intervention to reduce the increasing risk of social isolation (SI) and develop companionship in the older adult population. The COVID-19 pandemic provided a unique opportunity to examine the psychosocial intervention of befriending via technology use as a primary form of contactless socialization.

Objective: The study aims to explore the effectiveness of the befriending intervention through a contactless, intergenerational service-learning project on older adult emotions, especially boredom and loneliness as the key attributes of SI, and on students’ attitude toward companionship.

Methods: During the months of January to April 2022, undergraduate students enrolled in a health administration course with a special focus on culture were asked to be involved in a contactless, intergenerational service-learning project (n=46). In this study, contactless intervention meant communication using the telephone and apps such as FaceTime and Zoom. Students were paired with older adults to have at least a 30-minute weekly conversation, for 8 weeks, via telephone or an internet-based app such as FaceTime. Students were asked to write a half-page diary after each interaction and a 1-page reflection at the end of the fourth week and at the end of the service-learning project. At the completion of the project, the researchers also surveyed the older adults to assess the impact of the project using a 5-item open-ended questionnaire. Following a heuristic approach and content analysis, student artifacts (110,970 words; 118-page, single-spaced Microsoft Word document) and the older adult surveys were analyzed using MAXQDA, (VERBI GmbH). Qualitative data were extracted to assess the impact of service learning on SI by measuring the attributes of boredom and loneliness among 46 older adults. Students’ attitudes toward companionship were also assessed using data from their diaries and reflections.

Results: Overall, three major constructs were identified: (1) meaningful engagement, defined as feeling safe, having increased confidence, and having reduced boredom; (2) internal motivation to participate in the weekly interaction, defined as discussion about daily life experience, level of happiness, and ability to exert personal control over the situation; and (3) intergenerational befriending, defined as perceived benefits from the friendly nature of the interaction, ability to comfortably connect with students, and positive feeling and attitude toward the student.

Conclusions: The contactless, intergenerational befriending intervention reduced boredom and loneliness among older adults and enhanced positive attitude and confidence among university students. Students helped older adults to develop digital skills for the use of apps and social media. Older adults showed interest in the intervention and shared their daily life experiences with the students, which helped to reduce the gap between generations. Findings indicate the effectiveness of an intergenerational service-learning intervention on SI reduction and increased positive attitude among college students.
intergenerational befriending; social isolation; boredom; contactless socialization; service learning; internal motivation; mobile phone

**Introduction**

**Background**

The growing surge of the aging population has shone a spotlight on social isolation (SI). With increasing age, an increasing proportion of older adults experience chronic boredom and feelings of loneliness [1]. In addition, COVID-19 confinement and quarantine disturbed their lifestyle behaviors, making the aging population susceptible to high levels of stress and depression. SI is exacerbated when people are living alone, especially vulnerable older adults, experiencing the loss of family or friends, having comorbid or chronic illnesses, and facing sensory impairments. Even in residential settings and older adults’ living facilities, many older adults struggle to maintain their autonomy, self-determination needs [2], and competence as they are often disconnected from the broad community or dependent on unskilled nursing staff who lack social competence and training [2,3]. An intergenerational service-learning approach to befriending these individuals is an effective intervention to enhance their psychosocial well-being [4]. This approach demonstrates the importance of using reflective writing assignments to help university students deconstruct and reconstruct images, beliefs, and paradigms about older adults [5]. In addition, it is where the young individuals relate to older adults to discuss or share their life stories or day-to-day experiences to alleviate boredom, an attribute of SI [6].

Although SI is a global epidemic implying the absence of meaningful human relations and social connectedness, befriending is a psychosocial tonic to building social relationships and intergenerational friendships [7] that might reduce boredom and depression and significantly enhance the quality of life among older adults [8] and youth. The feeling of companionship is augmented when an intergenerational conversation is conducted purposefully, thereby negating the exacerbating consequences of SI such as boredom or even the feeling of loneliness. These negative subjective experiences result from discrepancies between an individual’s desires and perceptions about the quality of social relationships [9].

Intergenerational conversations are a gateway to enduring socialization for older adults and students [9]. The quality of life is enhanced as both generations engage in purposeful activities in a safe and friendly environment [8] to the satisfaction of each other’s psychosocial needs and well-being. Long-lasting intergenerational conversations can create a bonding between the older adult and the student, and it enhances the level of patience and builds social competence in youth, which are essential attributes to become skilled health care professionals in long-term care [10]. The social support provided by students can foster emotional health for the isolated, frail, older adult who is no longer at the center of a network of friends and acquaintances. Older adults with social connections have a high quality of life under social support [11], are motivated to work and engage in social activities, and have an inclusive attitude that promotes a healthy lifestyle [12]. Hence, they have low risk of conditions such as depression, additive behaviors, or dementia [13]. With social connectedness, even in living alone situations, these individuals have life satisfaction, with high cognitive stability [10,14].

**Facets of SI**

SI is one of the most disruptive transformations that exacerbate the quality of life, affecting human behavior, perceptions, and experiences [10,15]. It is a condition that leaves one with feelings of boredom and loneliness that are detrimental to physical and mental health [7]. Loneliness among older adults is the subjective, distressed feeling of being alone or separated, whereas SI is the objective physical separation from the community [14,16]. In addition, boredom is an emotion that often triggers negative thoughts such as self-harm and anxiety, and it is profound when the surrounding environment is mostly empty [17]. Hence, boredom, compounded by the feeling of loneliness over a prolonged period, may result in depression [18], which is a mental disorder that can be controlled in a timely way by identification of the needs of individuals and restoration of successful attention in meaningful activities such as intergenerational conversations. More than 56 million adults aged ≥65 years live in the United States, accounting for approximately 16.9% of the nation’s population [19]. In this population, >7 million (13%) are socially isolated. Of those 7 million individuals, 1.3 million are severely socially isolated [5,7]. The crucial drivers of boredom [18,20], namely, lack of recreational opportunities; limited personal contacts; immobility; prescribed home office; and, especially, COVID-19 quarantine and isolation have dissuaded the everyday lives of older adults, triggering anxiety and monotony [4,16,21].

Boredom across the population increased significantly owing to the COVID-19 pandemic containment measures [22], and so did the behavioral intention to find information, access services, and connect socially [23]. For older adults, there is evident increase of the benefits of technology interventions for social connectedness [10,24]. This may have helped the contactless, intergenerational, service-learning intervention to become especially effective, facilitating the feeling of companionship and social support that is felt bilaterally and actively by both parties [25].

**Intergenerational Service Learning**

Intergenerational service learning is an experiential learning, which is a course-based, credit-bearing, educational experience in which students participate in an organized, service-learning activity that meets identified community needs. Students reflect about the service activity to gain further understanding of a course content [26]. With an intent to develop a sense of companionship or friendship between both parties, the
intergenerational service-learning intervention brings the socially isolated older adults back on the periphery of social activities as they engage in conversations with college students [27]. Recurrent interactive sessions are useful to enhance the level of subjective well-being among older adults and connect them to the social network [28]. Alternatively, the service-learning sessions provide a broad appreciation of discipline and an enhanced sense of civic responsibility to the students while connecting them in purposeful activity with community members [29]. Requesting students to write diaries and reflections regarding their interactions with older adult adults is a key component of service learning [25,29]. Although it is less likely to see older individuals readily embracing service-learning sessions involving students or even social media sites when compared with young adults, adoption rates for contactless interactions with students among individuals aged ≥65 years have approximately doubled in the United States in the past 4 years [14,30].

Intergenerational Befriending Approach
This study used an approach that aims to bolster a long-lasting, genuine connection or relation between generations, especially college students engaged in service learning and older adults who are socially isolated [31]. This connection is based on shared experiences of daily living, which is reflected in the recurrent interaction based on the multicultural project. Although the intergenerational connection may develop over a contactless platform, it fosters a sentiment of compassion and empathy through focused interactions that add purpose and engagement for participants. The befriending idea is based on reciprocity and what benefits the older adult in a relationship [32]. The interactions are mostly based around topics that are meaningful to the older adult [10,33] and benefit the student to understand the biases of ageism. The idea is to make the older adult feel valued and cared for in a relation that is free from any kind of service delivery, obligation, or family ties [26,34].

Contactless Socialization
The quality of life comprises components such as health; well-being; peaceful existence; living in harmony; social engagement; life satisfaction; and keeping oneself busy with hobbies, volunteer service, or work [8,35,36]. The older adults were subjected to a harsher reality during the pandemic than younger adults as old age was affirmed as a risk for COVID-19 complications [37]. This contention promoted contactless socialization through contactless service-learning sessions via SMS text message, FaceTime, and emails that connected the older adults to wide social networks [38]. Although user confidence remained as a dominant issue when using technology such as a smartphone or tablet for interaction needs [39], the willingness and interest of older adults let students train them on the use of technology. The training also facilitated the use of mobile health and telemedicine among older adults, which greatly enhanced their health-related quality of life [40]. Hence, technology-based, contactless, service-learning sessions provided safe interaction for both parties, fostering social support and technology skill augmentation for the older adults [41]. Simple telephonic calls were used to build intergenerational friendships [13,38] connecting student helpers with their clients through personal life experience, interpersonal interaction, collaboration, and understanding [42].

Study Purpose
The study aimed to explore the effectiveness of the befriending intervention through a contactless, intergenerational, service-learning project on older adult emotions, especially boredom and loneliness as the key attributes of SI, and on students’ attitude toward companionship. Our approach was to forge an alliance and create an affective bond between the young student and the older adult to facilitate the possibility of friendship as they shared their life experiences and students engaged in technology training activities for the older adults regarding the use of apps and social media sites. The relationship that develops through befriending is seen as central to the experience while hypothesizing the facts that befriending would foster psychosocial well-being among older adults and that students’ attitudes toward older adults would become more positive throughout the service-learning course.

Methods
Study Setting and Recruitment
The research design was based on qualitative data analysis [43]. Data were collected through the survey of older adults and dialogues included in journal entries submitted by 46 undergraduate students enrolled in a Health Administration course. This core or required course had a special emphasis on cultural competency and diversity. The study followed 46 older adults, aged between 64 and 82 years, via intergenerational service-learning sessions during the months of January to April 2022. The older adults were recruited from residential facilities in Good Samaritan Society, Denton, Texas; Schertz Senior Living, Texas; Knowles Home, Nashville, Tennessee; Aguadilla Seniors, Puerto Rico; and Guadalajara Senior Center, Mexico. We selected these locations because these are the largest centers for older adults known to the researchers, and we asked the older adults in these centers to invite other participants known to them to participate in the study, to expand the sample size.

The inclusion criteria for older adult participants consisted of the following: those who (1) were interested in socialization activities, (2) were aged ≥65 years, (3) could read English, and (4) were willing to participate in the study. The exclusion criteria were the following: older adults who (1) were aged <65 years, (2) were already engaged in ≥1 socialization activity, and (3) did not pass the “attention check” in a meeting with the researcher conducted before the intervention to assess their interests and identify careless respondents, thereby improving the data quality. The inclusion criteria of student participants were the following: all students enrolled in the health administration course irrespective of age or involvement in socialization activity. No exclusion criteria for students were determined.

Data Collection Procedures
The service-learning sessions are an essential component of the undergraduate Health Administration course that has culture as
a major topic. The course introduces undergraduate students to the historical and cultural development of health care in contemporary American society. During the months of January to April 2022 a total of 46 students were paired with 46 older adults from the abovementioned communities. Once a week, the students communicated with their assigned older adult partners and engaged in an unscripted conversation for at least half an hour, for a total of 8 weeks. Of the 46 pairs, 29 (63%) pairs engaged in telephonic conversations and 17 (37%) pairs engaged in internet-based conversations via casual calling app, such as FaceTime or Zoom. The risk of COVID-19 contamination restricted all possibilities of in-person meetings.

As part of service learning, students were required to create an artifact diary to document each conversation. Students were also asked to write a 1-page diary after each interaction with their older adult partners and 2 reflections. On the basis of the duration of the project, students were expected to write 8 diaries and 2 reflections, once at the end of their fourth interaction and then again at the end of their eighth interaction. The conversations were recorded by the students with their smartphones or technology used for internet-based interaction such as Zoom and then transcribed by the students.

As part of the study, older adults were surveyed (paper-based, 5-item open-ended questionnaire) by the researchers to determine their interest in conversation to reduce SI and evaluate the effectiveness of the project. This 1-time survey was mailed with return envelopes to the older adults at the end of the eighth interaction with students. Survey questions were open ended, so that older adults could write their answers. The students’ transcribed conversations, diaries, and reflections and the older adults’ answers to the open-ended questions in the survey were all used in the qualitative data analysis.

The open-ended survey questions for the older adults included the following: (1) Did you find the conversation interesting? If your answer is yes, please write a few lines what was interesting in the conversation. (2) Would you like to participate in our project again next semester?

The whole idea of the project was to provide a useful framework for befriending that may facilitate and create a meaningful bond between the young student and the older adult [44,45]. All questions were composed from previously validated survey instruments and contextualized for use in this study (refer to sources in Table 1). To assess older adult emotions, especially boredom and loneliness as the key attributes of SI, students focused on conversations (refer to definition in Table 1) based on life satisfaction and digital skill training in the context of meaningful engagement [46,47], internal motivation to participate in the intervention for both older adults and students [47-49], and human feelings in the context of intergenerational befriending or companionship [46,49,50]. Thus, in Table 1, we have presented 3 major constructs: meaningful engagement, internal motivation to participate, and intergenerational befriending.

Table 1. Study constructs, definitions, and sources.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Definition</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaningful engagement</td>
<td>• Frequency of socially interactive activities</td>
<td>Questions about satisfaction with life, contextualized from the studies by Diener et al [46] and Gierveld and Tilburg [47]</td>
</tr>
<tr>
<td></td>
<td>• Feeling of reduced boredom</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Cohesive interaction (confidence and safety)</td>
<td></td>
</tr>
<tr>
<td>Internal motivation to participate</td>
<td>• Discussion about daily life experience</td>
<td>Questions about social isolation and motivation, contextualized from the studies by Kozma and Stones [48], Gierveld and Tilburg [47], and Russell [49]</td>
</tr>
<tr>
<td></td>
<td>• Level of happiness</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ability to exert personal control over the situation</td>
<td></td>
</tr>
<tr>
<td>Intergenerational befriending</td>
<td>• Perceived benefits from the friendly nature of interaction that develops mutual trust</td>
<td>Questions about human feelings, contextualized from the studies by Diener et al [46], Russell [49], and Golden et al [50]</td>
</tr>
<tr>
<td></td>
<td>• Ability of the older adults to comfortably connect with students</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Positive feeling and attitude of the students</td>
<td></td>
</tr>
</tbody>
</table>

Ethical Considerations

This contactless, intergenerational, befriending interventional study was approved by the institutional review board (protocol number 2022-7046) of the Texas State University. Written informed consent was obtained from all participants (students and older adults) before the intervention. Participants had the option to exit the intervention after reading the informed consent information or to provide consent to participate in the study. The confidentiality of the participants was properly protected during the intervention and data analysis. The study data were fully deidentified. All records pertaining to the intervention were securely protected in the university database with protected passwords, which were only accessible by the researchers. As this study was not grant funded, participants (undergraduate students and older adults) were not compensated for participating in this study.

Data Analysis

The older adults’ survey was mailed to the 46 older adult participants, and 21 (46%) returned the completed survey. Students’ diaries and reflections were collected at the end of the intervention (46/46, 100%), and all the files were used for data analysis. The qualitative data used in the study to extract the constructs and subconstructs were obtained from the 21 older adults’ surveys and 46 students’ diaries and reflections. Our analytic data included all the 110,970 words in a 118-page, single-spaced Microsoft Word document.
Following a heuristic approach [51] and content analysis, the data were coded using the qualitative text analysis software, MAXQDA by numbering each line of the dialogue [52]. The “Advanced Coding Method” in MAXQDA Standard was used for content analysis with major keywords that were allocated to data segments. The “Lexical Search” function located the keywords in all the text that define the 3 major constructs as identified in Table 1. Once the keywords of befriending, SI, boredom, socialization, engagement, and motivation were identified, the thematic coding of the relevant texts was performed using MAXQDA’s visual tool, “One Code Model.” We explored the frequency of words and terms used in the sources and analyzed their semantic contexts in a quantitative way. A differentiated word frequency analysis was performed using the “MAXDictio Module” [52]. The data were then classified into several groups to reveal trends and patterns of response to each question in the survey and in each topic of conversation between the older adults and students from the student diaries and reflections. Topically similar codes were grouped together and then narrowed by code segments using the option “Subcode Statistics” in the context menu of the “Code System” of MAXQDA. For example, the code “intergenerational befriending” was further subcoded as “comfortability,” “positive emotions,” and “perceived benefits.” Finally, using the MAXQDA function “Analysis Summary Grid,” thematic compilations were presented in “Participant Comment Tables.”

The use of heuristic inquiry [51,52] helped us to discover the nature of social phenomena, especially the intergenerational bonding that developed between participants, as we systematically coded the data. Owing to the complex nature of the emerging themes related to befriending and self-expression, heuristic inquiry appeared to be the most convenient method of sense making for this study. The heuristic depictions of the artifacts involved a synthesis of intuition and tacit understanding of researchers [53]. This understanding characterized the idea of befriending experience. The concepts that emerged from the study included relationship building, empathy, social interactions, and capacity for additional relational networks outside the family.

Results

In the older population of 46 adults, of which 19 (41%) were men and 27 (59%) were women, 3 main constructs were identified: intergenerational befriending, meaningful engagement, and internal motivation to participate in the program. According to the older adult participants, the perceived benefits of the friendly nature of intergenerational interaction were mainly the heightened feelings of comfort and reduced boredom. For both the student and older adult participants, trust and compassion were noticed in the initial stages, turned to compassion, trust, respect, empathy, honesty, and warmth in the later stages, which are the most essential attributes of friendship. The discussions about daily life experience led to emotive bonding and friendship. For the students, the befriending experience (specified 201 times) enhanced the feeling of comfort (specified 70 times) and positive emotions (59 times). Both the students and older adult participants looked forward to the sessions and considered it as a reason to get up in the morning.

The intergenerational service-learning intervention mainly covered two activities: (1) discussions about daily life experience to reduce the gap between generations and (2) digital training to help older adults to use social media sites and apps such as Facebook or Uber. The main attributes of SI targeted in this study, namely, boredom and loneliness, were minimized to some extent with the befriending approach as people felt safe and happy as they interacted with the students. Importantly, older adults were extremely interested in the technology training from students regarding how to use social media sites and apps such as Facebook and Uber. Tables 3 and 4 reflect the confirmatory statements of the older adults and students toward the valuable “befriending” component that engaged older adult participants substantially with the undergraduate college students in digital training and motivated them to share their life experiences safely. The attributes of hesitation and fear, which were noticed in the initial stages, turned to compassion, trust, respect, empathy, honesty, and warmth in the later stages, which are the most essential attributes of friendship.

The positive attitude and the dedication of the students toward the project helped the older adults to find social support and feel safe. On most occasions, activities were based upon the interest of the older adults. The opportunities for social interaction were possible as students were proactive and flexible with time and the needs of their older adult partners. The results revealed the importance of students’ positivity to support the older adults and the increased level of motivation for spontaneous interactions. The students created a friendly environment, or that of companionship, where the older adults had confidence that the discussions were appropriate based on their choices and preferences.

Another key finding was that older people define boredom and comfort differently and that there is a difference between the students’ perceptions and older adults’ perceptions related to interactions to alleviate boredom. Although befriending through
these sessions cannot compensate for the loss of an attachment figure, such as a close friend, spouse, or a significant other, which is common in old age, people with poor social skills are likely to have trouble in developing and maintaining relationships. Negative perceptions about age and aging, at societal and individual levels, have adverse effects on older adults’ health and well-being. Ageism, which means negative attitudes toward older adults, or unrealistic expectations about the intergenerational sessions can leave both the older adults and the young individuals with unmet social needs, resulting in increased boredom. Hence, the befriending approach must be understood from the individual’s subjective point of view.

Table 2. Constructs and subconstructs identified from the intergenerational service-learning intervention.

<table>
<thead>
<tr>
<th>Name of the constructs and subconstructs</th>
<th>Number of times specified by older adults (n=786), n (%)</th>
<th>Number of times specified by students (n=557), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intergenerational befriending</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comfortability</td>
<td>256 (32.6)</td>
<td>179 (32.1)</td>
</tr>
<tr>
<td>Positive emotions</td>
<td>79 (10.1)</td>
<td>70 (12.6)</td>
</tr>
<tr>
<td>Perceived benefits</td>
<td>80 (10.2)</td>
<td>59 (10.6)</td>
</tr>
<tr>
<td>Perception of happiness</td>
<td>97 (12.3)</td>
<td>50 (8.9)</td>
</tr>
<tr>
<td>Internal motivation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling of happiness</td>
<td>277 (35.2)</td>
<td>177 (31.8)</td>
</tr>
<tr>
<td>Discussion about daily life experiences</td>
<td>96 (12.2)</td>
<td>91 (16.3)</td>
</tr>
<tr>
<td>Meaningful engagement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feeling safe</td>
<td>253 (32.2)</td>
<td>201 (36.1)</td>
</tr>
<tr>
<td>Increased confidence</td>
<td>58 (7.4)</td>
<td>49 (8.8)</td>
</tr>
<tr>
<td>Reduced boredom</td>
<td>98 (12.5)</td>
<td>97 (17.4)</td>
</tr>
<tr>
<td>Reduced boredom</td>
<td>97 (12.3)</td>
<td>55 (9.9)</td>
</tr>
</tbody>
</table>

Table 3. Comments of older adults about the key constructs.

<table>
<thead>
<tr>
<th>Key construct</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaningful engagement</td>
<td>• “Like to talk about you and the multicultural sensitivity project in the Zumba classes three times a week, and ballet classes twice a week.” [Participant 2]</td>
</tr>
<tr>
<td></td>
<td>• “Feel glad, I seem to have more control on anger unlike before, we have opportunities for social interaction and there is less counterproductive behavior.” [Participant 45]</td>
</tr>
<tr>
<td></td>
<td>• “She plays for the university’s basketball team. Since the university is so close to where I live, I get to attend all her games and is eternally grateful for this.” [Participant 5]</td>
</tr>
<tr>
<td></td>
<td>• “The student group hosts a couple of social events every day, so there is always an event to attend and enjoy. Some of these events include- bingo, bible studies, stretching class, and physical fitness classes as well.” [Participant 42]</td>
</tr>
<tr>
<td></td>
<td>• “She is physically very active, I enjoyed our zoom discussion last week, me and my husband, engaging in either yoga, meditation, or going on 30-minute walks.” [Participant 16]</td>
</tr>
<tr>
<td>Internal motivation to participate</td>
<td>• “Family reunion was over; it was a good day of our weekly conversation and i-phone training.” [Participant 39]</td>
</tr>
<tr>
<td></td>
<td>• “Grateful to be able to wake up each morning and do the things that I enjoy, like art and seeing family and talking to my student friend over the phone.” [Participant 3]</td>
</tr>
<tr>
<td></td>
<td>• “Like to talk to you about school, do you enjoy the simple pleasures in life like cooking?” [Participant 28]</td>
</tr>
<tr>
<td>Intergenerational approach to befriending</td>
<td>• “Now look forward to our weekly conversations and app training sessions.” [Participant 40]</td>
</tr>
<tr>
<td></td>
<td>• “I don’t get to see my grandchildren as often as I would like, so having a person of the younger generation to talk to weekly reminds me a lot of them.” [Participant 25]</td>
</tr>
</tbody>
</table>
Table 4. Comments of students about the key constructs.

<table>
<thead>
<tr>
<th>Key construct</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Meaningful engagement                | • “At first the project was challenging, towards the end I felt relaxed, as I came to know the older adult partner more closely.” [Participant 9]  
|                                      | • “She explained it so well, I never thought before that smoothie making could be so simple.” [Participant 11] |
| Internal motivation to participate   | • “She loves everything about her life. Our hobbies are the same. She is blessed to be healthy at her age, and she wouldn’t change a thing.” [Participant 12]  
|                                      | • “My grandmother used to tell me similar things, to be able to wake up each morning and enjoy breakfast with family.” [Participant 1] |
| Intergenerational approach to befriending | • “Feels nice to be in an interesting conversation and write diaries.” [Participant 46]  
|                                      | • “So happy to help someone who sees her grandchild in me.” [Participant 45] |

Discussion

Summary

The befriending intervention delivered through the service-learning sessions promoted reciprocity and social support that enhanced the quality of life by fostering positive emotions and reduced boredom and loneliness. The intergenerational service-learning program fostered discussions about daily life experience that reduced the gap between generations and promoted companionship [54]. The proactiveness and dedication of the students for their project helped the older adults to find social support and feel safe and created a bonding between the generations with long-lasting friendships.

Meaningful Engagement

Through this specific intergenerational service-learning experiment, we found that most older adults try to keep themselves occupied through personal hobbies; talking to family over the phone; and occasionally, even work. Meaningful engagement was depicted well through the comments of older adult participant 42 who was happy and well engaged in programs and activities administered by the university such as telephone-based befriending and contactless luncheons with students. For these participants, social interaction through the service-learning program enhanced connectivity and reduced the risk of losing the motivation to maintain an active and healthy lifestyle. The conversations reduced the feelings of boredom and isolation that adversely affected their emotional health. These sessions helped the older adult to reengage in community networks and participate freely in cohesive neighborhood environments with other members of the older adult living facility. Older adult participant 2 found these sessions to be meaningful and interesting, explained the importance of these sessions to friends and community, and constantly spoke about the interactions in the Zumba and ballet classes 2 to 3 times a week. The intergenerational service-learning conversations toward the end of the project showed a heightened level of comfort for both the older adults and the youths and made the older adults feel that they are cared for while in conversation. Hence, they were not bored in their conversation with the students.

Internal Motivation to Participate

For the older adults, there is the need of a companion to share life experiences [55], which is profound when separated from the family owing to relocation, death of family members, or retirement [56]. The befriending activity reduced the boundaries between the 2 people involved, and the student befriender was asked to undertake tasks that the older adult wanted to do, such as using technology or having contactless visits during family reunion. This created a sense of internal motivation for older adult 40 to participate in the interactions. As students helped the older adults to interact with their family and share life stories, these interactive sessions fostered companionship that was of interest to the older adult and a reason for the student to find meaning and value in the conversations. Following the comments of older adult participant 25 and student participant 45, we found discussions about spouses to children and grandchildren and, for some, even to great grandchildren. All familial connections showed the importance of social support and motivated both parties in conversation on a day-to-day basis. Student participant 12 prayed for the well-being of his older adult partner and felt happy as he discovered a common hobby to talk about. Older adult participant 39 described her spouse returning from rehabilitation and mentioned that a welcome home party with her spouse’s family resulted in her having a good day. Older adult participant 5 revealed that she is eternally grateful to be able to attend her befriender’s basketball games; the joy and motivation to be able to socially connect seemed to have a positive effect on her well-being.

Intergenerational Approach to Befriending

Although several of the older adults in this study had various outlets to socialize such as hobbies, calling a friend, and family visits outside the intergenerational service-learning sessions, they still expressed appreciation toward the students for their time and companionship, which fulfilled the idea of befriending that we aimed for in this study. Student 45 noted that the older adult was nervous at the beginning but eventually looked forward to their weekly conversations, further mentioning that having someone from the young generation to talk to reminded her about her grandchildren whom she does not get to see often. In another instance, student participant 46 would write down events worth mentioning every week to discuss during the call. The overall evaluation of the intergenerational conversation was positive; older adult 40 wrote in the survey that she valued the ability to talk, listen, and share information with another...
human being whom she could trust and rely on. The befriending approach within these sessions appeared to expose the similarities between 2 different generations as the weeks passed. For example, student 28 revealed that he felt lucky to share the same religious and spiritual beliefs as the older adult partner. Similarly, student 42 stated that she and her older adult partner shared the same type of church. Most of the student’s reflections revealed the idea of reliability, compatibility, intimacy, and reciprocity in the conversations, which had a positive effect on promoting friendship and reducing boredom and loneliness for both parties. In addition, the program improved the youth’s social skills and assumptions related to one’s privacy and safety.

Limitations
This study has some limitations. Several students reported being unable to communicate with their partners at the beginning of the project owing to difficulty with older adults’ digital literacy, inability to answer phone calls or read SMS text messages, or inability to use apps such as Zoom or FaceTime. This ruled out the possibility of Zoom focus groups. In-person focus groups were also not possible owing to COVID-19 socialization restrictions. In some cases, there was dropout of older adults from the project owing to sickness or other health conditions. Although the authors clearly explained to the older adults that the service-learning project was completely contactless, some of them still preferred face-to-face interaction, which was not possible because of COVID-19 restrictions. Furthermore, student notes may not be free from bias. The findings from this study have limited generalizability because we used a sample of students from 1 course, and the results may not be applicable in other study settings. Variables such as overall health, stress, stigma, comorbid conditions, socioeconomic status, and discrimination were not considered while assessing older adult surveys or student artifacts in the study.

Conclusions
The need to feel important and be included in a large social group is vital regardless of age. The need is perhaps more during the later stages of life when one has more physical and mental challenges and is confined in living situations with very limited social networks. The effects of service-learning sessions on boredom and loneliness among older adults promoted interest in intergenerational practice to reduce SI and improved the feelings of companionship among both students and older adults as a long-term outcome of the intervention. Students identified the intergenerational component as a highlight of the health administration course, which increased their awareness of gerontological issues and knowledge about working with aging populations. The intervention enhanced students’ attitude toward companionship and gerontological practices via intergenerational befriending opportunities and digital training sessions. As students developed an interest in service-learning programs through frequent interactions with older adults, the social interaction component in the program amplified social connectivity via contactless interactions, phone conversations, and digital training sessions for older adults. Consequently, it reduced the risk of losing the motivation to maintain a self-image and an active and healthy lifestyle among older adults. This study was designed for periodic or postpandemic follow-up of the participants experiencing SI. Our results showed that older adults wished to maintain their social connectedness and access to technology and digital networks. Driven by a desire to communicate and access information that stemmed from the need to connect to friends, family, or other internet-based services such as mobile health or social media, older adults showed appreciation and gratitude toward the youth for taking an interest in their lives and, often, would look forward to their weekly conversations.

Acknowledgments
This study received no external funding. The authors are grateful to all the older adult study participants for their time and support in the study. The authors are grateful to Dr Stan Ingman for his assistance with the recruitment phase. The authors are also grateful to Dr Gayle Prybutok for helping with qualitative data analysis.

Data Availability
The institutional review board of this study does not allow data sharing because the data protection statement is an integral part of our institutional review board approval. Therefore, our data will not be made available.

Authors’ Contributions
KS was involved in conceptualization, methodology, validation, formal analysis, investigation, resources, data curation, writing, and original draft preparation. ZR and NL wrote and prepared the original draft. All other authors reviewed and edited the manuscript.

Conflicts of Interest
None declared.

References


17. LaMonica HM, Davenport TA, Roberts AE, Hickie IB. Understanding technology preferences and information requirements for health technology devices designed to improve and maintain the mental health and well-being of older adults: participatory design study. JMIR Aging 2021 Jan 06;4(1):e21461 [FREE Full text] [doi: 10.2196/21461] [Medline: 33404509]


38. Chopik WJ. The benefits of social technology use among older adults are mediated by reduced loneliness. Cyberpsychol Behav Soc Netw 2019 Sep;19(9):551-556 [FREE Full text] [doi: 10.1089/cyber.2016.0151] [Medline: 27541746]


41. Meinert E, Milne-Ives M, Surodina S, Lam C. Agile requirements engineering and software planning for a digital health platform to engage the effects of isolation caused by social distancing: case study. JIMIR Public Health Surveill 2020 May 06;6(2):e19297 [FREE Full text] [doi: 10.2196/19297] [Medline: 32348293]


Abbreviations

SI: social isolation

©Keya Sen, Nida Laheji, Zo Ramamonjiarvelo, Cecil Renick, Randall Osborne, Brad Beauvais. Originally published in JMIR Aging (https://aging.jmir.org), 30.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.

https://aging.jmir.org/2024/1/e47908 JMIR Aging 2024 | vol. 7 | e47908 | p.57

Edited by J Wang; submitted 05.04.23; peer-reviewed by J Neils-Strunjas, S Santini; comments to author 03.08.23; revised version received 23.08.23; accepted 28.12.23; published 30.01.24.

Please cite as:
Sen K, Laheji N, Ramamonjiarvelo Z, Renick C, Osborne R, Beauvais B
Examining the Effect of Contactless Intergenerational Befriending Intervention on Social Isolation Among Older Adults and Students’ Attitude Toward Companionship: Content Analysis
JMIR Aging 2024;7:e47908
URL: https://aging.jmir.org/2024/1/e47908
doi: 10.2196/47908
PMID: 38175944
Determinants of Implementing an Information and Communication Technology Tool for Social Interaction Among Older People: Qualitative Content Analysis of Social Services Personnel Perspectives

Johanna Fritz, PhD; Petra von Heideken Wågert, PhD; Annelie K Gusdal, PhD; Rose-Marie Johansson-Pajala, PhD; Caroline Eklund, PhD
School of Health, Care and Social Welfare, Mälardalen University, Eskilstuna/Västerås, Sweden
*all authors contributed equally

Corresponding Author:
Johanna Fritz, PhD
School of Health, Care and Social Welfare
Mälardalen University
Box 883
Eskilstuna/Västerås, 721 23
Sweden
Phone: 46 21101503
Email: johanna.fritz@mdu.se

Abstract

Background: Older people are particularly vulnerable to social isolation and loneliness, which can lead to ill-health, both mentally and physically. Information and communication technology (ICT) can supplement health and social care and improve health among the vulnerable, older adult population. When ICT is used specifically for communication with others, it is associated with reduced loneliness in older populations. Research is sparse on how the implementation of ICT, used specifically for communication among older people in social services, can be performed. It is recommended to consider the determinants of implementation, that is, barriers to and facilitators of implementation. Determinants related to older people using ICT tools are reported in several studies. To the best of our knowledge, studies investigating the determinants related to the social services perspective are lacking.

Objective: This study aims to explore the determinants of implementing the Fik® room, a new, co-designed, and research-based ICT tool for social interaction among older people, from a social services personnel perspective.

Methods: This study used an exploratory, qualitative design. An ICT tool called the Fik® room was tested in an intervention study conducted in 2021 in 2 medium-sized municipalities in Sweden. Informants in this study were municipal social services personnel with experience of implementing this specific ICT tool in social services. We conducted a participatory workshop consisting of 2 parts, with 9 informants divided into 2 groups. We analyzed the data using qualitative content analysis with an inductive approach.

Results: The results included 7 categories of determinants for implementing the ICT tool. Being able to introduce the ICT tool in an appropriate manner concerns the personnel’s options for introducing and supporting the ICT tool, including their competencies in using digital equipment. Organizational structure concerns a structure for communication within the organization. Leadership concerns engagement and enthusiasm as driving forces for implementation. The digital maturity of the social services personnel concerns the personnel’s skills and attitudes toward using digital equipment. Resources concern time and money. IT support concerns accessibility, and legal liability concerns possibilities to fulfill legal responsibilities.

Conclusions: The results show that implementation involves an entire organization at varying degrees. Regardless of how much each level within the organization comes into direct contact with the ICT tool, all levels need to be involved to create the necessary conditions for successful implementation. The prerequisites for the implementation of an ICT tool will probably change depending on the digital maturity of future generations. As this study only included 9 informants, the results should be handled with care. The study was performed during the COVID-19 pandemic, which has probably affected the results.
KEYWORDS
information and communication technology; implementation; determinants; social isolation; loneliness; organization; digitalization; facilitators; barriers; older people

Introduction

Older people are especially vulnerable to social isolation and loneliness, particularly because they are exposed to risk factors such as living alone and experiencing chronic illness [1]. There is robust evidence linking loneliness and isolation with physical decline; morbidity; increased mortality; and cognitive and mental health problems, such as depression and dementia; and increased risk of suicide [2-7]. Several studies point toward an increase in loneliness during the COVID-19 pandemic with stay-at-home orders and recommendations for social distancing [8-11], and the impact is particularly severe among people aged ≥80 years [12]. Information and communication technology (ICT) can supplement health and social care and improve health in the vulnerable, older population [4,13]. ICT is a part of welfare technology, which in one way or another, improves the lives of those who need it. When technology is used specifically for communication with others, it is associated with reduced loneliness [4,13-16] and increased well-being and life satisfaction [4] in older populations. Technology can improve social connectedness among older adults. The specific effectiveness rates favor ICT and videoconferencing [13]. Chen and Schultz [15] identified 4 important mechanisms for reducing social isolation using ICT: staying connected to other people, such as family and friends; gaining social support; participating in interesting activities; and boosting self-confidence. Studies show that ICT can support and maintain the social relationships and healthy and independent lives of older people at the individual level and should be prioritized as an early and preventive intervention in social services [17]. However, the use of ICT has been shown to decrease after 6 months of use [15,16]. Only few studies have investigated how the implementation of web-based social activities in social services can be performed. Thus, there is a need for sustainable, structured, and well-planned solutions for the implementation of ICT in social services.

To make ICT useful for older people, social services must consider the determinants for its implementation, that is, barriers to and facilitators of implementation. Determinants related to older people are reported in several studies. A literature review including 59 papers identified determinants related to this specific population’s adoption of technology, such as perceived usefulness, potential benefits, user friendliness, ease of learning, perceived costs and savings, knowledge about existence, availability in the market, technical support, social support, perceived emotional and psychological benefits, and relevance with their previous experiences [18]. Other identified determinants related to older people are gaps in ICT literacy, fear of making mistakes when learning the ICT tool [19], privacy concerns, technical difficulties, lack of user-friendly options designed specifically for an older population, and lack of experience in using technology [20,21]. However, implementation of ICT does not depend on the older people alone. Social services personnel have an important role in introducing ICT to older people and to support its use [22]. Thus, ways of working to introduce and support the use of ICT among older people need to be implemented in the social services and other services they offer. To the best of our knowledge, determinants of implementing an ICT tool for social interaction among older people related to the social services personnel perspective are lacking.

The readiness and maturity to adopt digitalization and new ways of working vary among the social services provided by Swedish municipalities, for example, in residential care services and home care services. Few older people receive access to welfare technology services despite the benefits [23]. It is a large step for an organization to move from a limited project to implementation in their organization. Konttila et al [24] identified determinants of importance for digitalization in health care but not specifically for the care of older people, related to professionals’ knowledge, skills, attitudes, and experiences and organizational and collegial support. One of the proposals from the Swedish Ministry of Health and Social Affairs [23] is that studies are needed to implement models for welfare technology in social services. A systematic review focused on facilitators and barriers that influence the implementation of welfare technology for older people, from the perspectives of older individuals, people with disabilities, informal caregivers, health and care personnel, organizations, infrastructure, and technology [25]. Overall, 6 themes of determinants were identified: capacity, attitudes and values, health, expectations of effects, shared decision-making, and identity and lifestyle. These determinants are within different levels in an organization and are consistent with other determinant frameworks for implementation [26,27]. However, most of the included papers in the systematic reviews of determinants for digitalization and implementation of welfare technology for older people [24,25] involved various types of technology, such as technology for smart homes, mobile devices in medicine and public health, self-care, medication, and surveillance systems, whereas ICT used specifically for communication among older people was not included. This study aimed to explore the determinants of implementing the Fik® room, a new, co-designed, and research-based ICT tool for social interaction among older people, from a social services personnel perspective.

Methods

Design

This study used an exploratory, qualitative design [28,29]. An exploratory design is appropriate for conducting studies in a field that is relatively underexplored and hence, an inductive approach was adopted [30].
The Fik@ room: An ICT Tool for Social Interaction Among Older People

The Fik@ room is a research-based ICT tool, in the form of a web platform for safe web-based social interaction, created and developed by researchers in coproduction with older people aged ≥60 years, municipal health and social care personnel, and an IT company, based on focus group interviews and workshops. The content and design of the Fik@ room was based on the needs and wishes expressed by older people [31]. The older people who participated in the development process in 2019 and 2020 expressed the importance of knowing that others visiting the Fik@ room were equal, that is, experiencing loneliness, and access to the Fik@ room was protected by an individual password. This log-in procedure contributes to increase the safety and the feeling thereof for the older people. All older people with access to the Fik@ room have received personal log-ins from a contact person from the municipality. People who received log-ins were older people who had experienced loneliness or social isolation [31]. The Fik@ room consists of digital coffee tables with seating for up to 4 people at each table (Figure 1). The older people can start conversations regarding topics of their own choice using video, voice, or chat. They can also post messages to each other on a bulletin board. In the Fik@ room, older people can meet new friends and socialize as a way of forming routines in their everyday lives. The Fik@ room focuses on meeting peers (people in the same situation) and offers the opportunity to meet and discuss subjects related to users’ interests. This foundation improves the quality of conversations in the Fik@ room and increases independence, participation, accessibility, and options for users to form their own social interactions as part of their everyday lives, which is associated with better quality of life for older people [17]. The Fik@ room is available on Google Play Store and Apple App Store (for iPad devices) but will not be available for logging in without permission from an authorized gatekeeper such as personnel in the municipality.

Figure 1. Illustration of a digital coffee table within the Fik@ room. The conversation theme of this table is sports, and in this example, only 1 person is attending.

Informants and Settings

The ICT tool was tested in an intervention study performed in May and June 2021 in 2 medium-sized municipalities in the middle of Sweden. The 2 municipalities were chosen because they are coproduction partners of the University and were both involved in the development of the Fik@ room. Informants in this study were social services personnel, with experiences of the intervention study and thus the implementation of this specific ICT tool in social services: municipal social services managers and social services personnel who recommended the ICT tool to older people in the intervention study (Table 1). The informants were chosen using a purposive sampling technique to capture different perspectives and experiences from the stakeholders in the implementation process, which provided the
process with experience-related information and theory-based knowledge. In total, 9 informants from the 2 municipalities agreed to participate, who were basically all the people involved in the implementation process. Municipality 1 had an ongoing digitalization project running in parallel with the intervention study. This means that the municipality was appointed by The Swedish Association of Local Authorities and Regions as 1 of 10 model municipalities, which would function as a model for the digitalization of care for older people. These 10 municipalities received extra financial support to enable time for knowledge dissemination. Together with The Swedish Association of Local Authorities and Regions, they support other municipalities with knowledge about digital services and welfare technology. Municipality 2 was not involved in the parallel, ongoing digitalization project.

Table 1. Characteristics of the informants and settings.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Municipality 1 (n=6)</th>
<th>Municipality 2 (n=3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex of the informants, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>2 (33)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Female</td>
<td>4 (67)</td>
<td>3 (100)</td>
</tr>
<tr>
<td>Profession of the informants, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manager</td>
<td>2 (33)</td>
<td>1 (33)</td>
</tr>
<tr>
<td>Developer</td>
<td>1 (17)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Occupational therapist</td>
<td>0 (0)</td>
<td>1 (33)</td>
</tr>
<tr>
<td>Technology supporter^a</td>
<td>2 (33)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Guide for older people^b</td>
<td>1 (17)</td>
<td>1 (33)</td>
</tr>
<tr>
<td>Ongoing digitalization project</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

^aTechnology supporters were IT experts, employed in the municipality, who prepared the iPad devices for the older people in the intervention study by installing SIM cards and connecting them to the network and installing the Fik@ room app.

^bGuides for older people were social services personnel with experience in IT, employed in the municipality, who worked at the meeting places for older people and supported the older individuals regarding the use of the iPad and the Fik@ room app.

Data Collection

Participatory workshops with 2 researchers acting as workshop leaders provided a valuable opportunity to learn together and discuss several perspectives. The informants both generated and analyzed data through a structured process for data collection and analysis that includes a combination of individual and group activities, inspired by the effect modifier assessment (EMA) method [32]. The EMA method consists of workshops and subsequent analyses. The workshop leaders facilitate the collection of information on past significant events; in this study, the event in focus was the introduction of the Fik@ room to older people. The method facilitates a combination of individual and group activities, which implies that all experiences from each informant is used and then developed in a group discussion. In this study, semiquantitative estimations were not used because all the determinants (barriers and facilitators) were considered important regardless of how many people had experienced them. The workshop guide (Multimedia Appendix 1) could be considered as a semistructured interview guide, but the workshop informants interactively influenced the interview guide by deciding what events and in what way the events are discussed. Each informant participated in 1 or 2 participatory workshops in August and September 2021, in groups of 3 to 6 participants. Each workshop began with casual conversation to help the informants feel at ease and more comfortable in the setting. The researchers served as workshop leaders to encourage a flow of discussion. To create a comfortable environment for the informants, the workshops were conducted separately for each municipality, ensuring that all the informants within a workshop were familiar with one another. The participatory workshop consisted of 2 parts. The first part focused on the determinants for implementing the ICT tool, and the workshop activities involved the identification of possibilities, obstacles, and challenges that were experienced during the intervention study. The first workshop leader (JF) asked the informants to individually note the possibilities, obstacles, and challenges on a paper in front of them and, thereafter, facilitated a group discussion regarding the same questions. The second workshop leader (CE) asked clarifying questions during the workshop and summarized the discussion at the end of the workshop. The second part started with a review of the first part, followed by the same individual and group processes as the first part, focusing on scenarios for the development of methods to support the implementation of the ICT tool. The workshop activities were regarding how to overcome the obstacles and reinforce the possibilities identified in the first part of the participatory workshop.

The workshops were performed using a web-based video communication tool, Microsoft Teams (Microsoft Corp). Municipality 1 performed the 2 parts of the workshop in separate sessions (2 hours each), and municipality 2 performed both the parts in the same session (2 hours). The informants’ professions were requested at the beginning of the first workshop. The workshops were video recorded using the Microsoft Teams video platform.

https://aging.jmir.org/2024/1/e43999

JMIR Aging 2024 | vol. 7 | e43999 | p.61
(page number not for citation purposes)
Data Analysis

The data were analyzed using qualitative content analysis with an inductive approach [33]. The recordings were transcribed verbatim. The text was read several times for familiarization. Meaning units related to the determinants for implementing the ICT tool for social interaction among older people from the perspectives of the social services personnel were identified, coded, and grouped into subcategories and categories according to similarities. Examples of the abstraction of categories from meaning units are shown in Table 2. The identification of meaning units and categorization were performed by the first author. To validate the analysis, part of the categorization was also performed by 2 other authors (CE and PvHW) separately. In addition, the analysis was regularly discussed and validated among all the authors during the process to achieve consensus.

Table 2. Examples of the abstraction process: meaning units, codes, subcategories, and categories.

<table>
<thead>
<tr>
<th>Meaning unit</th>
<th>Code</th>
<th>Subcategory</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>...That they [the personnel] have a login so when they are there, they [personnel and older person] can talk to each other at a table [within the ICT tool] just like we did.</td>
<td>Be able to log in to the ICT tool</td>
<td>Be able to show the tool</td>
<td>Be able to introduce the ICT tool in an appropriate manner</td>
</tr>
<tr>
<td>If it had been as usual [before the pandemic] then we would have done the same at home visits, and it would have been much easier because then you can show this leaflet, talk about it, and sell it in a better way, so it would have been much better.</td>
<td>Easy to show in person</td>
<td>Be able to show the tool</td>
<td>Be able to introduce the ICT tool in an appropriate manner</td>
</tr>
<tr>
<td>The pandemic, of course, because it has not been possible to visit people. They have not wanted to let us in, and not even homecare staff have been able to visit some people. They have declined home care and arranged help in another way because they do not want to expose themselves to the coronavirus.</td>
<td>The pandemic hindered in-person visits</td>
<td>Be able to show the tool</td>
<td>Be able to introduce the ICT tool in an appropriate manner</td>
</tr>
</tbody>
</table>

*ICT: information and communication technology.

Ethical Considerations

This study was conducted in accordance with the Declaration of Helsinki [34] and Swedish Ethical Review Act [35]. However, according to the act, ethics approval by the Swedish Ethical Review Authority is not needed when, for example, sensitive personal data are not collected (ie, when interviewing staff to determine how they perform their work), as in this study. However, the intervention study (in which older people were study participants) has obtained ethics approval by the Swedish Ethical Review Authority (Dno 2020-06640). Participation was voluntary, and all informants provided their verbal consent after receiving verbal and written information. Furthermore, they were allowed to withdraw at any time without consequences.

Results

Overview

The results included 7 categories of determinants for implementing the ICT tool: be able to introduce the ICT tool in an appropriate manner, organizational structure, leadership, digital maturity of the social services personnel, resources, IT support, and legal liability (Textbox 1). The results are presented using the categories as headings, and the subcategories are italicized in the text.
Textbox 1. Categories and subcategories of determinants for implementing the information and communication technology (ICT) tool for social interaction among older people.

<table>
<thead>
<tr>
<th>Be able to introduce the ICT tool in an appropriate manner</th>
</tr>
</thead>
<tbody>
<tr>
<td>• To tailor the information</td>
</tr>
<tr>
<td>• Be able to show the ICT tool</td>
</tr>
<tr>
<td>• To let the user practice</td>
</tr>
<tr>
<td>• The personnel need to have knowledge about the ICT tool</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organizational structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>• A system for effective communication</td>
</tr>
<tr>
<td>• Clear roles</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Leadership</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The manager’s engagement</td>
</tr>
<tr>
<td>• Enthusiasts</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The digital maturity of the social services personnel</th>
</tr>
</thead>
<tbody>
<tr>
<td>• The personnel’s digital skills</td>
</tr>
<tr>
<td>• The attitudes among the personnel varied</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Resources</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Costs</td>
</tr>
<tr>
<td>• Time</td>
</tr>
<tr>
<td>• Access to transport</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>IT support</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Accessibility</td>
</tr>
<tr>
<td>• Follow-up system</td>
</tr>
</tbody>
</table>

Legal liability (no subcategories)

Be Able to Introduce the ICT Tool in an Appropriate Manner

Older people need to be informed about the existence of the ICT tool. Reaching all potential users with information about the tool was a challenge that the informants did not know how to address. They found that written information sent via mail reached many potential users and was easy to distribute. However, when introducing the ICT tool to an older person, the informants found it important to tailor the information, for example, the amount of information that the older person was able to receive, according to their knowledge and attitude toward ICT. Knowing the person and meeting in person facilitated tailoring compared with written information and web-based meetings:

> I have to know how to structure the conversation with the person I am calling, and I also have to do that when I call the person because I have to hear what status the person has, how should I handle the person, i.e., how should I structure my conversation so that I establish good communication with the person. I must choose my conversational tone mode, how I present it, and how I tell it, a lot is about structuring the conversation and I have to do that immediately when I get in touch with the person. [Informant 4]

When introducing the ICT tool, the informants found it important to be able to show the tool. To enable this, the personnel also needed to have the prerequisites to log in to the ICT tool, which was not always a matter of course. The personnel also needed to have access to their own account, and the program needed to be installed in their digital equipment. To reduce older people’s fear of digital tools, the informants found it important to let the users practice using digital tools in a playful manner. Getting acquainted with digital equipment, such as tablets or computers, through playing games, watching movies, or reading newspapers can improve the ease of use of other digital tools such as the ICT tool. Some meeting places for older people offered these practices and integrated digital tools into their daily services. During the COVID-19 pandemic, in-person visits were not always possible but were considered a prerequisite for showing and practicing the ICT tool:

> To dare to use the tablet...to use it for something they are familiar with, read the newspaper or whatever it may be, as a first step. It may not be the ICT tool that is the first step, but it may be the next step when they have learned to use the tablet. A game can be a little
To be able to provide information regarding the ICT tool, personnel need to have knowledge about it. The informants perceived the written information regarding the ICT tool as useful for understanding the purpose of the tool. However, it was difficult to inform older people about a tool without knowing how it looked or how it worked. Having both seen and tried the tool made the personnel more confident when informing older people about it:

It was great that we got to try the ICT tool first, because it also makes it easier when you are out with the user to show them how it works because you actually know what the picture looks like, how the sound sounds, what happens with the sound if we sit too close to each other. So, I thought it was great, you need to try it first. [Informant 6]

**Organizational Structure**

When implementing the ICT tool, the informants perceived it as important to have a system for effective communication within the organization. A secure electronic communication system regarding personal data worked appropriately for communication among different parts of the organization regarding lists of potential users, who to call, who would visit whom, information needed for home visits, and so on. After the personnel had visited an older person, they used the same system for feedback about the visit and to document whether further support was needed. The informants also highlighted the importance of feedback among different levels within the organization, such as to the management team by whom new decisions could be made:

Lists were created so that we had a structure on which we had called, who would go to whom, all the information needed to make a home visit and even be able to write when you had been there, and it was a completed project...So, it was a very good structure in the lists. [Informant 3]

Spontaneously, I would say that some structure will be required for the recruitment of these [older] people. [Informant 2]

The informants expressed that clear roles facilitated the implementation of the ICT tool. Clarity about whom to ask regarding a special issue, whom to ask for support, and so on makes the work more efficient. Similarly, the person who holds the role knows what duties come with the role. They experienced that the implementation was facilitated if the selected personnel performed the introduction, skills training, and support of the ICT tool:

I thought a bit about this regarding whom to contact and so on. It should be incredibly clear, both for our users but for the employees as well. They should not have to think “who are we to contact to get support for this?” but there must be somewhere very, very clear so it can be done quickly, so that you do not have to run around and look and waste time looking for who to contact. [Informant 7]

**Leadership**

The informants expressed that the manager’s engagement was important to the implementation, and it became obvious when engagement was lacking. A manager who was open to the ICT tool and interested in its implementation spread their engagement to the rest of the working group. According to the participants, another success factor was having enthusiasts or champions as leaders. These were selected personnel with a clear mandate within the organization who had extra knowledge about ICT, who worked actively regarding its implementation, and who regulated the entire process:

That it is someone who owns the question, who has the question on their table, who is the one who then ensures that it is followed up, and the continuity of the whole thing I think is very important. [Informant 6]

**The Digital Maturity of the Social Services Personnel**

The implementation of the ICT tool was affected by the social services personnel’s digital skills. According to the informants, some of the personnel did not know how digital equipment worked, such as a tablet, which hindered them from informing others about and supporting the use of the ICT tool:

The personnel couldn’t use a tablet either, you had to give them basic instructions on how to press the button to start it up. Of course, there were also those who were very talented. But you might think that it is only the older people that don’t have the skills, but it is actually the personnel too. [Informant 3]

The informants expressed that the attitudes among the social services personnel varied. Some were very interested, and the implementation was conducted smoothly. In other parts of the organization, the personnel did not even talk about the ICT tool. The informants thought that an ICT tool that can be used by both older people and personnel would create great interest among the personnel compared with a tool that can be used only by older people. For example, the personnel could conduct lectures or discussions about health-related subjects, such as diet and exercise, within the ICT tool. The informants thought that this digitalization investment had a positive impact on personnel’s attitudes toward digital tools, which provided synergies and paved the way for the implementation of this specific ICT tool:

Yes, but what if we can have a table where we can talk about health and diet, and those who want to can come in and hear, listen, or participate in discussions. [Informant 2]

**Resources**

Implementation requires resources of various types. In this case, the informants highlighted resources regarding costs, time, and access to transportation. They expressed concern about the costs that would be required to gain access to the tool and support. The implementation of the ICT tool would be at the expense of something else. According to the informants, a
payment model based on different fees, depending on what is included, may facilitate implementation:

If you start from the scenario that it is the municipalities that in some way buy a license or something similar for the ICT tool, then the payment model itself could be for a lower amount if the municipality itself, so to speak, moderates what is said and not, and perhaps a higher amount then if it is a company who would be responsible for it. [Informant 5]

Implementation takes time, and some of the most time-consuming parts, as mentioned by the informants, were sending and following up on information letters, delivering tablets, creating log-ins, and showing users what to do. The informants felt that time had been allocated at different levels (e.g., to key individuals within the digitalization investment area of the model municipality):

It does not matter what we are going to do, time is required. And if you have decided to make a change like this, you should be aware that time is needed. So, it is obvious that it is something that really needs to be considered if it would be implemented somewhere else as well. That you actually make sure you have that time and resources, it costs to implement something, but in the end, it can generate so much more. [Informant 7]

Access to transportation is a prerequisite for home visits. The informants experienced that there are always cars available as a means of transportation for home visits.

**IT Support**

The accessibility of the support, both for the users and personnel, was perceived as important for implementing the ICT tool. Contact information needed to be available, and it was preferable if all support could be reached using the same contact method (e.g., the same phone number or email address). In addition, time needs to be allocated for support. According to the informants, support not only involved direct contact with the user but also involved communication with and recruiting feedback between the supporters and the personnel. Support could be provided through various forums to increase accessibility, such as during home visits, at meeting places for older people, or at the public library. Support could also be provided via other digital media, which hindered accessibility, as the requested support was sometimes related to the difficulties in handling the digital equipment itself and, thus, also the digital support. During the COVID-19 pandemic, the accessibility of support was particularly limited, as digital support could not be received owing to some users’ lack of skills, and the older people did not accept home visits owing to the risk of spreading the infection:

What needs to be strengthened is time, it is the key that we have talked about. It was also what was difficult, but we also saw that what still worked well was when there was allocated time for various steps: time to be at home with the user to practice and provide support and follow-ups, but also time for communication with personnel and recurring feedback: to ask how it works out. So, I think that resources and time really need to be strengthened, then you have all the prerequisites to succeed. [Informant 6]

The personnel lacked a follow-up system for support. They expressed a need to be able to follow up regarding whether the user had used the ICT tool to facilitate implementation. It was not possible to assess whether the older person understood how to use the ICT tool after a short introduction. The informants suggested that it should be possible to obtain information about the number of log-ins on the ICT tool to be able to follow up with users who have few log-ins:

It would be interesting to get feedback on if these users have not been in at all. Then you could have maybe called them and asked: How are you? Do you want more help? [Informant 9]

**Legal Liability**

The informants expressed doubts about whether the municipalities complied with the legislation if the ICT tool was offered by the municipality and used inappropriately, such as the use of racist statements. There was a concern that the municipality cannot guarantee that nothing inappropriate is said within the ICT tool without some form of supervisory function. Therefore, it was proposed that a moderator of the ICT tool could perform that function:

If the municipality buys it, we stand as some form of guarantor, we also have a responsibility not to release that freely, but to have some form of moderating function that can support what is said. Because racist statements may be used, for example, and then we also have a responsibility to take care of it. [Informant 5]

**Discussion**

**Principal Findings**

The results included 7 categories of determinants for implementing the ICT tool. Being able to introduce the ICT tool in an appropriate manner concerns the personnel’s options for introducing and supporting the ICT tool, including their competencies in using digital equipment. Organizational structure concerns a structure for communication within the organization. Leadership concerns engagement and enthusiasm as driving forces for implementation. The digital maturity of the social services personnel concerns the personnel’s skills and attitudes toward using digital equipment. Resources concern time and money. IT support concerns accessibility, and legal liability concerns possibilities to fulfill legal responsibilities.

**Comparison With Previous Studies**

Determinants of implementing ICT tools related to older people have been reported in several studies and systematic reviews [18-21]. The novelty of our study lies in its knowledge about the determinants of implementing a new, co-designed, and research-based web platform, customized specifically for older adults, from a social services personnel perspective. The
determinants identified in this study are concretized to increase the understanding of specific factors that influence the implementation of an ICT tool for social interaction among older people. Damschroder et al [27] and Flottorp et al [26] highlighted several domains of determinants, including factors related to the implemented intervention itself (in this case, the ICT tool) and patient factors. These 2 domains are not presented in our results but are reported in a related article [22]. The social services personnel’s view about the determinants identified in this study were mainly related to the interaction between the personnel and older individual, and organizational factors. Our findings differ from those of other studies describing older people’s views about determinants, which were mainly related to their own capacity, attitudes, and health-related benefits and the usefulness and ease of learning the technical tool [18-22]. However, the interaction between the professional and the older individual relates to strategies for supporting older people to overcome barriers mentioned as determinants by the older people themselves. Thus, although the determinants mentioned by social services personnel and older people differ, they are logically interconnected. The 7 categories of determinants reported in this study correspond to the following domains reported in the paper by Flottorp et al [26]: individual health professional factors; professional interactions; incentives and resources; capacity for organizational change; and social, political, and legal factors. Some of the factors within these domains were not mentioned as determinants by the informants in our study, such as the continuing education system, assistance (external) for organizational change, contracts, and political stability. As they were not mentioned by the informants, we interpreted them to be of less importance in this specific case.

Most of the identified determinants in our study were related to contextual factors at different levels, which confirms that contextual determinants play an important role in implementation [36]. Nilsen and Bernhardsson [36] highlighted contextual factors as determinants at the micro (interaction between the professional and patient), meso (the organization), and macro (influences from the wide environment) levels. The micro-meso-macro framework for analysis is a useful way of understanding the determinants of implementation, as implementation is a multilevel phenomenon [37]. In this study, the contextual determinants mostly involved the micro (be able to introduce the ICT tool in an appropriate manner) and meso levels (organizational structure, leadership, resources, and IT support). Only 1 category was identified at the macro level (legal liability). The Lancet and Financial Times Commission on Governing Health Futures 2030 [38] recommends interventions at the macro level to facilitate the digitalization of health and social care to achieve future health and well-being. From the perspective of the personnel, the impact of the identified determinants likely differs. As macrolevel determinants were not mentioned by the informants to the same extent as microlevel and mesolevel determinants, it could be assumed that macrolevel determinants were not perceived to have as great an impact as micro- and mesolevel determinants on the implementation of the ICT tool.

A category of determinants that we identified, the digital maturity of the social services personnel, was related to individual health professional factors according to the checklist by Flottorp et al [26] and, in particular, knowledge, skills, and cognition. A lack of digital competence has been identified across all professions within social services in Sweden, and the development of the personnel’s competence is stated to be a success factor when implementing welfare technology in social services [23]. Konttila et al [24] recommend that learning how to use technical devices should be integrated into the personnel’s daily work by providing education and sufficient time for learning. In previous studies, knowledge and skills primarily focused on digital and technical aspects. However, our results in the category, be able to introduce the ICT tool in an appropriate manner, also emphasize the importance of the personnel’s pedagogical knowledge and skills in teaching older people how to use an ICT tool. This introduces additional demands on the personnel that must be considered when implementing an ICT tool in social services.

Our results are consistent with the barriers to and facilitators of the implementation of welfare technology identified by Zander et al [25]. All our identified determinants correspond to the themes reported by Zander et al [25] regarding capacity, attitudes, and values. In addition, we identified enthusiasts, as part of the leadership category, as an important driving force for implementation, which was not explicitly mentioned by Zander et al [25]. However, similarity can be seen with the theme of participation, as Zander et al [25] discussed the importance of involvement in the development, decision-making, and implementation processes as a facilitator of implementation. A theme of determinants that Zander et al [25] identified but were not identified in our results was expectations. Expectations were seen as a barrier to the implementation of welfare technology and were related to fear that the technology would affect the quality of care, threats to professional identity, and fear of losing jobs. It is possible that the ICT tool in our study was not perceived as a threat against the personnel’s professional role, as it did not directly affect the quality of care or replace the care provider. It is also important to remember that the technology used in the literature review by Zander et al [25] did not include ICT used specifically for communication among older people, which can explain the differences in the results.

According to our results, enthusiasts seemed to be an important facilitating determinant for implementation. Enthusiasts can also be described as champions or local opinion leaders, depending on whether they are appointed by the management or considered informal, educationally influential leaders appointed by peers [39]. In our results, we interpret enthusiasts more consistently with the definition of champions. However, it remains unclear whether the enthusiasts only function through managerial status and process or also function through social influence, such as an opinion leader. To support implementation, the evidence for the role of local opinion leaders is more robust than that for champions, and it seems that involvement of local opinion leaders is an effective implementation strategy [39,40]. To understand the impact mechanisms of enthusiasts, the role and significance of enthusiasts need further clarification.

One of the 2 municipalities included was a model municipality for the digitalization of care for older people, which contributed
to an important difference between these municipalities. The informants from the model municipality expressed the determinants by describing their own experiences as facilitators (eg, their experiences regarding how a system for effective communication among personnel facilitated the implementation of the ICT tool). The other municipality talked about the same determinant as a barrier, that is, the lack of a system for effective communication. It was obvious that the digitalization project positively influenced the implementation of the ICT tool, which also confirms the importance of determinants related to the organization. Although the ICT tool was supposed to be easy to use for older people [31], it is still important to have an organizational structure to support the older people in using the ICT tool. Previous studies highlight the importance of a shared vision within the organization for the implementation and involvement of leadership [25]. These determinants were perceived by the informants in the model municipality. They also experienced other facilitators identified in previous studies, such as a system for communication, clearly defined roles, enthusiasts, access to IT support, and resources that facilitated the implementation [24,25]. Digital maturity appears to be great in the model municipality, which also affected the implementation of the ICT tool.

Implementation is more likely to be successful if implementation strategies are chosen based on an assessment of determinants (facilitators and barriers) of implementation [40]. However, when selecting the implementation strategies, consideration must also be given to the effectiveness of different strategies based on well-designed studies and systematic reviews, the phases—implementation or maintenance—of the actual implementation process [41], and the theoretical underpinnings of the implementation [42,43]. Our findings can contribute to increasing the understanding of the complexity of implementing an ICT tool for social interaction among older people in municipal settings and guide the choice of implementation strategies.

Several studies have explored the determinants of using ICT tools, often in relation to older people. ICT tools can mean different types of technology that are used in health and social care and used by older people outside health and social care. Most ICT tools studied are not designed specifically for older people [18-21,23-25]. In this study, determinants are explored in relation to the Fik® room, a specific ICT tool for social interaction among older people, developed for and in coproduction with older people. As ICT tools involve large variety, it can hinder the transferability of our results. Therefore, it is important that the readers themselves are aware of the type of ICT tools that are studied.

**Strengths and Limitations**

As determinants for improving professional practice have been identified at different levels [26,41], we wanted to include the informants involved in the implementation of the ICT tool in different ways. The combination of informants working with leadership and informants working directly with the older people in different ways enabled a comprehensive understanding of the determinants of the implementation of an ICT tool for social interaction among older people in municipal settings. It would have been desirable to have more informants from municipality 2. Although the study only included 9 informants, these informants had specific experiences pertinent to the study’s aim by being involved in the intervention study that introduced the Fik® room to older people. Taken together with the specific phenomenon in question and a narrowly defined objective, the data were considered to have sufficient information power [44].

The use of participatory workshops inspired by the EMA method [32] contributed to a systematic approach to data collection and opportunities to learn from each other. The combination of individual and group activities meant that all experiences from each informant were used and developed in group discussions, which contributed to multifaceted and rich data. By including a second part of the workshop, focusing on how to overcome obstacles and reinforce the possibilities identified in the first part of the workshop, we were able to further use the informants’ thoughts and experiences.

Owing to the COVID-19 pandemic, the workshops were performed using the web-based video communication tool, Microsoft Teams. Although conducting qualitative studies over the internet facilitated the informants’ participation, it also involved some challenges [45]. A challenge was the informants’ familiarity with technological hardware and software. In this case, the informants had access to a reliable internet connection and a computer with a microphone and camera. The video communication tool, Microsoft Teams, was chosen because the informants were already familiar with this program because it was used in their organizations. All the informants (9/9, 100%) used digital communication tools in their daily work and were familiar with how to communicate using them. Another challenge is that web-based interviews seem to generate short responses and less contextual information [45], which could also apply to web-based workshops. Compared with in-person workshops, web-based groups need to be small, and 4 to 6 participants are recommended [45]. Therefore, we included a maximum of 6 informants in each workshop group. It might be a limitation that the informants in the workshops knew each other. A power imbalance might have occurred between managers and employees, which might have limited the issues that were raised for discussion during the workshops. However, the opinion of the 2 researchers participating in the workshops was that the informants spoke freely regarding the barriers and facilitators of implementing the ICT tool. The fact that the implementation of the ICT tool was a project might have stimulated the informants to speak more freely than they might have done if the ICT tool was implemented as compulsory working task.

An exploratory design was considered appropriate for conducting research in this relatively underexplored field and thus, we adopted an inductive approach [30]. The choice of design was made to stimulate an open and creative discussion to enable the identification of new areas of determinants. In hindsight, a deductive approach would also have been possible to use because our results proved to be consistent with the checklist by Flottorp et al [26]. However, the use of the inductive approach contributed to a deep understanding of the determinants of implementing an ICT tool for social interaction among older people.
The findings present the determinants for the implementation of an ICT tool for social interaction among older people in municipal settings, but the determinants related to the ICT tool itself and older people are presented in a related article [22]. This division may complicate the possibility of obtaining an overall picture of the current determinants of importance for the implementation of the ICT tool. To make this easy for the reader, we refer to the related article by Gudsal et al [22] in the text.

The careful description of the data collection process and illustration of the findings with quotations increased the confirmability of the findings. Trustworthiness was strengthened through regular discussions among the authors during the analyses. The agreement between our findings and those of previous studies regarding the determinants of implementation in general [26] and implementation of welfare technology for older people in particular [25] increases the credibility of our findings, which is an important strength of our study and indicates wide transferability of the findings to the implementation of other ICT tools for older people in other contexts. However, the trustworthiness and transferability of the study results should be considered with caution because there were only 9 informants from 2 municipalities involved. Despite the small sample size, the study revealed important aspects to be considered when implementing ICT tools in municipality social services.

This study was conducted during the COVID-19 pandemic. The purpose was not to relate the results to the situation of older people specifically during the pandemic. However, the pandemic has probably affected the results in different ways. The problem of loneliness among older people increased during the COVID-19 pandemic [9-12], which increased the need for new solutions for communication among older people. This may have affected the attitude of both the older people and the personnel participating in this study toward ICT tools. The difficulties with meeting older people in person are also strongly associated with the restrictions during the pandemic and might have affected the results as the personnel did not have the prerequisites to meet older people and introduce the ICT tool (and the hardware) in the same way as without the pandemic and physical restrictions. Therefore, determinants with a direct connection to the pandemic are not as relevant during periods without a pandemic.

Conclusions
The ICT tool discussed in this study will be used by older people in their homes for the purpose of social interaction. Although the ICT tool is for use by older people in their homes, with limited involvement of others, the results show that the implementation involves an entire organization at different levels. Specifically, the following may be required: ability of the personnel to introduce and support the ICT tool, including their competencies in using digital equipment; structure for communication within an organization; leadership as a driving force; sufficient resources; and possibilities to fulfill legal responsibilities. Regardless of how much each level within the organization comes into direct contact with the ICT tool, all levels need to be involved to create the necessary conditions for successful implementation. The prerequisites for the implementation of an ICT tool will probably change depending on the digital maturity of future generations. As this study only included 9 informants, the results should be considered with caution. The study was performed during the COVID-19 pandemic, which has probably affected the results.

Acknowledgments
The authors wish to thank Ulrika Florin, PhD in innovation and design, for the valuable input regarding the design and prerequisites of this study. The authors also thank all the study informants for sharing their experiences. This study was supported by internal research funding from Mälardalen University, Eskilstuna and Västerås, Sweden.

Data Availability
The data sets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Authors’ Contributions
JF, CE, AKG, RMJP, and PvHW were involved in the methodology of the study. JF and CE conducted the workshops. JF conducted the content analysis and wrote the original draft. CE, AKG, RMJP, and PvHW supported the analysis and contributed to reviewing and editing the paper.

Conflicts of Interest
Coauthors CE, AKG, RMJP and PvHW own equity in the Fik® room, a company that might benefit from this study’s findings. The company was founded after the data collection of this study and aims to make the Fik® room available for its target audience. The first author (JF) led data analysis and manuscript preparation to ensure objectivity in the research.

Multimedia Appendix 1
Workshop guide.
[PDF File (Adobe PDF File), 110 KB - aging_v7i1e43999_app1.pdf]
References


Factors That Influence Successful Adoption of Real-Time Location Systems for Use in a Dementia Care Setting: Mixed Methods Study

Lynn Haslam-Larmer¹, RN(EC), PhD; Alisa Grigorovich¹,², PhD; Leia Shum¹, MASc; Andria Bianchi¹,³,⁴, PhD; Kristine Newman¹, RN, PhD; Andrea Iaboni¹,⁵, MD, DPhil; Josephine McMurray⁶, PhD

¹KITE Research Institute, Toronto Rehabilitation Institute, University Health Network, Ontario, ON, Canada
²Recreation and Leisure Studies, Brock University, St. Catherines, ON, Canada
³Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
⁴Centre for Clinical Ethics, Unity Health Toronto, Toronto, ON, Canada
⁵Department of Psychiatry, University of Toronto, Toronto, ON, Canada
⁶Lazaridis School of Business & Economics, Wilfrid Laurier University, Brantford, ON, Canada

Corresponding Author:
Josephine McMurray, PhD
Lazaridis School of Business & Economics
Wilfrid Laurier University
73 George St
Brantford, ON, N3T3Y3
Canada
Phone: 1 519 756 8228 ext 5649
Email: jmcmurray@wlu.ca

Abstract

Background: Technology has been identified as a potential solution to alleviate resource gaps and augment care delivery in dementia care settings such as hospitals, long-term care, and retirement homes. There has been an increasing interest in using real-time location systems (RTLS) across health care settings for older adults with dementia, specifically related to the ability to track a person’s movement and location.

Objective: In this study, we aimed to explore the factors that influence the adoption or nonadoption of an RTLS during its implementation in a specialized inpatient dementia unit in a tertiary care rehabilitation hospital.

Methods: The study included data from a brief quantitative survey and interviews from a convenience sample of frontline participants. Our deductive analysis of the interview used the 3 categories of the Fit Between Individuals, Task, and Technology framework as follows: individual and task, individual and technology, and task and technology. The purpose of using this framework was to assess the quality of the fit between technology attributes and an individual’s self-reported intentions to adopt RTLS technology.

Results: A total of 20 health care providers (HCPs) completed the survey, of which 16 (80%) participated in interviews. Coding and subsequent analysis identified 2 conceptual subthemes in the individual-task fit category, including the identification of the task and the perception that participants were missing at-risk patient events. The task-technology fit category consisted of 3 subthemes, including reorganization of the task, personal control in relation to the task, and efficiency or resource allocation. A total of 4 subthemes were identified in the individual-technology fit category, including privacy and personal agency, trust in the technology, user interfaces, and perceptions of increased safety.

Conclusions: By the end of the study, most of the unit’s HCPs were using the tablet app based on their perception of its usefulness, its alignment with their comfort level with technology, and its ability to help them perform job responsibilities. HCPs perceived that they were able to reduce patient search time dramatically, yet any improvements in care were noted to be implied, as this was not measured. There was limited anecdotal evidence of reduced patient risk or adverse events, but greater reported peace of mind for HCPs overseeing patients’ activity levels.

(JMIR Aging 2024;7:e45978) doi:10.2196/45978
KEYWORDS
remote sensing technologies; dementia; real-time location systems; Fit between Individuals, Tasks, and Technology framework; FITT framework; technology implementation

Introduction

Background
The increasing demands for efficiency and improvements in the delivery of care have prompted interest in the use of tracking technologies as a solution to monitor the movements of patients, providers, and equipment across health care settings, including hospitals and long-term care homes. As an example of this type of technology, real-time location systems (RTLSs) can be used to identify the physical location of people or assets across time and space [1,2]. RTLS technologies are local positioning systems that typically consist of a wireless device attached to objects or worn by people, with environmentally embedded reference points that receive wireless signals from these wireless devices. Software connected to this wireless network can provide continuous real-time visualizations of the location data on a facility map. RTLS installations have been studied in a wide variety of health care settings to monitor individuals’ movements (eg, residents or patients and health care providers [HCPs]) [3-5] and assets (eg, surgical equipment) [6] and, more recently, to collect health data [3,7-9]. These systems are particularly well suited for monitoring movement and reducing the risk of unaccompanied exit (eg, elopement); however, RTLS data have also been used to inform clinical decision-making processes and to monitor health status or the effects of an intervention [10]. Similar systems using wireless geolocation have already been successfully implemented in other sectors, such as insurance [11] and telecommunications [12], but present novel infrastructure and maintenance challenges, such as a lack of motivation for providers to engage due to a lack of understanding about the systems and poor communication about how they work. In contrast, enablers of RTLS adoption in acute care settings have prompted interest in the use of tracking technologies to improve the quality of residents’ lives or the workflow of HCPs. RTLS adoption in acute care settings has been slow, with concerns regarding provider privacy pitted against the goal of patient safety and efficiency [3].

Objectives
While many technology adoption models are in use [4] and have provided considerable insight into the adoption of digital technologies in the health care domain [5,6], to date, no study has specifically examined the implementation of an RTLS in a tertiary dementia care setting. To better demonstrate the potential barriers and enablers of RTLS adoption in a clinical setting, we used the Fit Between Individuals, Tasks, and Technology (FITT) framework developed by Ammenwerth et al [8]. The FITT framework suggests that the adoption of new technology in a clinical environment will depend on the fit between the attributes of the system or technology (eg, real-time location system) and the attributes of the clinical task (eg, degree of complexity and workload) [8]. The framework can be used retroactively [9] or prospectively to identify “deltas” between the expectations of a technology’s implementation in a clinical setting and its actual relevance in the workflows and to its users [15]. The FITT framework’s strength is its emphasis on the interaction between the individual and the task, with the understanding that harmonization will positively impact the implementation and adoption of novel technologies. In this study, we aimed to identify the factors that influence the adoption or nonadoption of an RTLS during its implementation in a secure inpatient unit for persons living with dementia in a rehabilitation hospital.

Methods

Study Design
This partially mixed methods concurrent study [16] included a brief preinterview survey followed by an in-depth semistructured interview. We chose this study design due to the exploratory nature of the research and the combination of quantitative and qualitative data that provides a deeper and more comprehensive understanding of the topic [17,18].

Study Setting
This study was conducted in a 20-bed secure inpatient dementia care unit of a large urban rehabilitation hospital in Ontario, Canada. Patients were admitted primarily from long-term care homes to receive specialized assessment and a personalized care plan to manage agitation, aggression, and other behaviors that interfere with the quality of life and safety of individuals living with dementia and their caregivers (eg, formal and informal). The unit comprised a team of interprofessional care providers (eg, nurses, physiotherapists, occupational therapists, recreation therapists, and geriatric psychiatrists) with expertise in addressing the range of physical, mental, and emotional challenges associated with dementia. As the unit was a secure unit, the team was familiar with using technology to help
monitor patients’ health status and movements. For instance, the unit used the WanderGuard elopement prevention system (Securitas Healthcare), which alerted providers if patients attempted to leave the unit unaccompanied.

Participants

Participants in this study were HCPs who regularly worked on the unit where the technology was implemented (eg, nursing aides, nurses, allied health professionals, and unit leadership). There were approximately 40 frontline staff members who interacted with patients, in addition to an advanced practice nurse and a unit manager. HCPs working on the unit were exposed to the RTLS and were individually able to decide to engage with the technology. The participants were recruited via an email distributed by the unit manager to all HCPs and by word of mouth during training sessions, team meetings, and daily report huddles.

RTLS Intervention

The implemented RTLS was a commercially packaged ultrawideband live monitoring system installed to locate and track patients on the unit. The hardware consisted of wall-mounted beacons that created a local Wi-Fi mesh and wearable tags that provided location data within the networked area. The wearable tags were fitted to patients on the unit as nonremovable bracelets. All participants’ substitute decision makers were provided with an opportunity to consent (or refuse) to have the patient wear the bracelet and to separately consent (or refuse) to have the location data collected and stored for research purposes (eg, development of clinical algorithms).

In the nursing station, a tablet app provided a view-only dashboard of the unit map and a live feed of patient locations. The tablet was locked to “kiosk mode” after log-in, thus preventing the app display from timing out. We held in-service training sessions to familiarize HCPs with the app’s layout and functions, including how to read the map and search for patients. Although the RTLS can be used in various ways (eg, nursing call bell), no other features were enabled in this study; the system was exclusively used to locate patients. The RTLS location data were stored on a secured and sectioned client cloud server approved by the Health Insurance Portability and Accountability Act.

Ethical Considerations

This study was approved by the University Health Network (UHN) Research Ethics Board (#20-6277) and the Wilfrid Laurier University Research Ethics Board (ID# 6767) and was conducted in accordance with the principles of the Declaration of Helsinki. The survey participants reviewed a web-based consent form, had an opportunity to contact the study team with any questions, and indicated their informed consent to participate through the UHN REDCap (Research Electronic Data Capture; Vanderbilt University) e-consenting process. Then, a personalized link was sent from the REDCap system via email for participants to complete the preinterview survey. Upon completion of the interview, HCPs received a gift card valued at CAD $20 (US $14.8).

Data Collection

Following installation, a research study team member (LS) held training sessions for HCPs to explain how the bracelets worked to track patient location and how to use the nursing station tablet and app. During the training sessions, HCPs were verbally informed that they would receive an email inviting them to participate in the study. Approximately 6 weeks after installation, HCPs were invited via email to participate in the study, which included a short survey followed by a more in-depth, semistructured interview. The survey (facilitated via the UHN-managed REDCap servers) took approximately 5 minutes to complete and included questions regarding their sociodemographic characteristics and attitudes toward the use of technology when caring for patients with dementia. These attitudinal questions asked HCPs regarding their perceptions of the following: (1) their comfort with technology generally and location-tracking technology in particular, (2) the appropriateness of using location-tracking technology on patients with dementia, (3) whether they planned to use the location-tracking system installed in their unit, and (4) the appropriateness of technology to monitor HCPs. Furthermore, staff were invited to provide free-text comments regarding why they planned to use or not use the technology. After completing the survey, the participants were contacted to participate in the semistructured phone interview. The interview comprised 12 primary questions (Multimedia Appendix 1), asking about their past experience (if any) using RTLS, expectations and goals of the system, user experiences with the system, and how they perceived the system to impact their care of patients. The interviews lasted approximately 1 hour and were audio recorded. The interviews continued until saturation was achieved.

Data Analysis

Qualitative analysis was conducted using NVivo (version 12; Lumivero). Anonymized interviews were transcribed verbatim. Transcripts were read by all the research team members (LHL, AG, JM, KN, and AB) to gain a comprehensive understanding of the content. Several team meetings were held to review the transcripts and develop a coding strategy. We used inductive and deductive analytic techniques, such as systematic coding and constant comparison to fit data with existing literature and to identify conceptual categories and insights [19]. After the first round of independent coding, the team met to ensure the optimal categorization of utterances within themes and subthemes. Each interview was coded by at least 2 team members. Group discussions and consensus were used to resolve any disagreements in coding.

Our deductive analysis of the interview data was informed by the FITT framework developed by Ammenwerth et al [8] and included the following categories: (1) individual and task, (2) individual and technology, and (3) task and technology. The FITT framework suggests that the adoption of new technology in a clinical environment will depend on the “fit” between the attributes of the individuals (eg, computer literacy), the attributes of the technology (eg, user interface and functionality), and clinical task attributes (eg, degree of complexity) [8]. The individual category represents not only the individuals using the technology but also groups of users (eg, HCPs) and considers...
constructs such as the physical settings in which they operate, thus representing any nontask phenomenon that may influence uptake and use of the technology. The technology category includes components of a digital application, such as hardware, software, and network infrastructure, or analog tools, such as paper-based care plans or manuals, used to complement a particular technology. The task category represents the clinical work and work-related processes that occur within a particular care setting.

**Results**

**Participant Demographics**

A total of 20 HCPs completed the preinterview survey. The survey’s primary purpose was to characterize the participant sample and allow the study team to determine whether the diverse range of HCPs on the unit was represented in the interview process. Of the 20 HCPs who completed the survey, 16 (80%) completed interviews, a sample size deemed sufficient given the exploratory nature of the single-site study and relative professional homogeneity of the sample population that allowed us to achieve saturation (Table 1) [20]. The participants who completed the survey and interview were predominantly female individuals (16/20, 80%), aged <50 years (16/20, 80%), and provided direct patient care (eg, nursing and support staff and allied health staff; 14/20, 70%).

<table>
<thead>
<tr>
<th>Measure and item</th>
<th>Surveys (n=20), n (%)</th>
<th>Interviews (n=16), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>16 (80)</td>
<td>14 (88)</td>
</tr>
<tr>
<td>Male</td>
<td>4 (20)</td>
<td>2 (12)</td>
</tr>
<tr>
<td><strong>Age group (y)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥25 to &lt;35</td>
<td>6 (30)</td>
<td>5 (31)</td>
</tr>
<tr>
<td>≥35 to &lt;50</td>
<td>10 (50)</td>
<td>8 (50)</td>
</tr>
<tr>
<td>≥50</td>
<td>2 (10)</td>
<td>2 (12)</td>
</tr>
<tr>
<td>Did not answer</td>
<td>2 (10)</td>
<td>1 (6)</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>College, other nonuniversity certificate, or diploma</td>
<td>7 (35)</td>
<td>5 (31)</td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>10 (50)</td>
<td>8 (50)</td>
</tr>
<tr>
<td>Master’s degree</td>
<td>3 (15)</td>
<td>3 (19)</td>
</tr>
<tr>
<td><strong>Occupation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Providing direct patient care (nursing and support staff)</td>
<td>14 (70)</td>
<td>10 (62)</td>
</tr>
<tr>
<td>Allied health (eg, OT(^a) or PTA(^b) and recreation therapy)</td>
<td>4 (20)</td>
<td>4 (25)</td>
</tr>
<tr>
<td>Leadership or administration</td>
<td>2 (10)</td>
<td>2 (12)</td>
</tr>
<tr>
<td><strong>Work experience (y)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤1</td>
<td>3 (15)</td>
<td>3 (19)</td>
</tr>
<tr>
<td>2 to 5</td>
<td>1 (5)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>6 to 10</td>
<td>3 (15)</td>
<td>2 (12)</td>
</tr>
<tr>
<td>&gt;10</td>
<td>11 (55)</td>
<td>10 (62)</td>
</tr>
<tr>
<td>Did not answer</td>
<td>2 (10)</td>
<td>1 (6)</td>
</tr>
<tr>
<td><strong>Current position (y)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤1</td>
<td>7 (35)</td>
<td>7 (44)</td>
</tr>
<tr>
<td>2 to 5</td>
<td>4 (20)</td>
<td>3 (19)</td>
</tr>
<tr>
<td>6 to 10</td>
<td>3 (15)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>&gt;10</td>
<td>4 (20)</td>
<td>4 (25)</td>
</tr>
<tr>
<td>Did not answer</td>
<td>2 (10)</td>
<td>1 (6)</td>
</tr>
</tbody>
</table>

\(^a\)OT: occupational therapist.

\(^b\)PTA: physiotherapy assistant.
All the participants who completed the survey agreed or somewhat agreed with the statements that they were “satisfied with their job,” were “confident in their ability to learn a new technology,” and were “familiar with location-tracking technology and how it works” and “agreed” or “somewhat agreed” that location-tracking technology was “acceptable when tracking the movements of persons living with dementia” (Table 2). There was less concordance of opinion regarding the use of location-tracking technology to monitor the movement of providers; 12 (60%) of the 20 respondents reported that its use would be “somewhat or completely unacceptable,” and 4 (20%) reported its use as “somewhat or completely acceptable” (Tables 3 and 4).

<table>
<thead>
<tr>
<th>Questions</th>
<th>Agree, n (%)</th>
<th>Somewhat agree, n (%)</th>
<th>Neither disagree nor agree, n (%)</th>
<th>Somewhat disagree, n (%)</th>
<th>Disagree, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>“I am satisfied with my job.”</td>
<td>15 (75)</td>
<td>5 (25)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>“Overall, I find technology is useful in my job.”</td>
<td>15 (79)</td>
<td>1 (5)</td>
<td>3 (16)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>“I feel confident in my ability to learn how to use new technology.”</td>
<td>17 (85)</td>
<td>3 (15)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>“I am familiar with location-monitoring technology and how it works.”</td>
<td>14 (70)</td>
<td>6 (30)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

Table 2. Survey question responses (N=20).

<table>
<thead>
<tr>
<th>Questions</th>
<th>Completely unacceptable, n (%)</th>
<th>Somewhat unacceptable, n (%)</th>
<th>Neither acceptable nor unacceptable, n (%)</th>
<th>Somewhat unacceptable, n (%)</th>
<th>Completely unacceptable, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>“In general, I believe that using location-monitoring technology to track the movements of persons with dementia is”</td>
<td>16 (80)</td>
<td>4 (20)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>“In general, I believe that using location-monitoring technology to track the movements of health care provider is”</td>
<td>1 (5)</td>
<td>3 (15)</td>
<td>4 (20)</td>
<td>7 (35)</td>
<td>5 (25)</td>
</tr>
</tbody>
</table>

Table 3. Survey question responses (N=20).

<table>
<thead>
<tr>
<th>Responses</th>
<th>Value, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Always</td>
<td>10 (50)</td>
</tr>
<tr>
<td>Often</td>
<td>5 (25)</td>
</tr>
<tr>
<td>Sometimes</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Rarely</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Never</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

Table 4. Survey question responses (N=20) to “I plan to use the location-monitoring technology (referred to as “RTLS”\textsuperscript{a}) during the pilot study at the SDU\textsuperscript{b}.”

\textsuperscript{a}RTLS: real-time location system.
\textsuperscript{b}SDU: special dementia unit.

### Interview Results

Interview data were analyzed and reported in 3 categories that correspond to the FITT framework [8]. The quality of the fit between these constructs depends on each of their characteristics and their alignment and the ability of management or the technology adoption team to influence the adaptation of the task, the technology, or the individual to improve the quality and success of the implementation.
the association between locating and monitoring patients for HCPs.

Characterization of the Task

In the first subtheme, participants identified the characteristics of the task, particularly the time-sensitive nature of some activities; for example, medication administration, which benefits from real-time monitoring, or conversely, the inability to locate a patient promptly when needed. One provider noted that the task of locating patients was more complicated with this specialized population as follows:

Because our patients are confused, and they have no concept of their own place or their own room, majority of them, they wander around the unit. [P5]

Legacy closed-circuit television cameras that streamed video in real time to monitors located at the nursing desk only allowed the visualization of public spaces (eg, hallways). A provider articulated the difficulty with locating a patient using this existing system, as follows:

[There’s no cameras in the patient rooms. Only in the hallway. Like [in] the public spaces...but not in the patient room. So, if you can’t see them in like the more public spaces, then you would actually have to get up and search every single room because the patients might not be in their room...They could be in someone else’s room, they could have fallen asleep on someone else’s bed, maybe they’re in someone else’s bathroom. [P21]

Furthermore, many HCPs shared that the task of locating patients manually made other tasks, such as medication administration, more complex:

Like if you wanted to give medication, [if] we couldn’t find, um, patients, we have to go and look for them in room by room. [P6]

Moreover, we conducted this study during the COVID-19 pandemic, during which patients who tested positive or were symptomatic of infection had to be isolated on the unit. While external doors were locked, most internal doors were not, particularly in patient rooms. The task of maintaining isolation for new patients, those who were exposed, or those who tested positive for COVID-19 in this unit was challenging, as they had difficulty remembering and understanding the need for isolation and physical distancing. HCPs had to monitor patients every 15 minutes to ensure that they stayed isolated and had to use a variety of strategies to encourage patients to remain in their rooms (eg, placing stop-strip door banners across the unit and stop signs on closed doors and sitting outside their rooms). Hallway cameras were perceived as inadequate for the task of supporting isolation, and providers felt that they did not allow for an anticipatory response when a patient was either moving from their bed or toward a restricted hallway or room.

In addition, this theme included the individuals’ need to track items other than human or physical assets in a dynamic space. One HCP stated that putting tags on nurses could help them locate each other when one of their patients required assistance. Another suggested that they would be useful for quantifying direct care hours provided by different providers.

Similarly, some participants suggested that the RTLS may be helpful for monitoring assets such as patient’s phones and walkers that are often misplaced. This is particularly important for assistive technologies, such as walkers, that support mobility and activities of daily living. One HCP noted as follows:

Someone who always uses their walker or, you know, always misplaces it...especially a lot of our patients now have, you know, like, personal items such as cell phones...and they’ve been misplaced and we’re always looking for it. [P8]

Especially with our...patient...group right now, a lot of them have their own cell phones. A lot of them, they put it in their pockets and a lot of the provider don’t really track where they are. So, when the patients are looking for it, we’re the ones who have to go around and look, you know, in the bag, in the laundry, in their closer. [P8]

Association Between Locating and Monitoring Patients for HCPs

The second individual-task subtheme accounted for an HCP’s perception that the act of locating or knowing the location of a patient on the unit was strongly associated with a better standard of care by helping to mitigate potentially “risky” events. One HCP noted as follows:

It gives me more stress if every time I keep on looking for my patient, I couldn’t get my things done right away, ‘cause I have to find them and always making sure that they’re safe. [P12]

Being unable to identify when one patient entered another patient’s room was a common concern noted by HCPs:

So once in a while, we do our rounds, but if a patient sneaked into somebody’s room, how would I know, until I really look at who is in bed? [P5]

We cannot find them because they wander around, you know, they...pace and...they are so intrusive, they go to other, uh, patients’ rooms. [P6]

A number of HCPs noted that “distractions” would often interrupt the process of physically locating a patient on the floor, for example:

Another patient come(s) up to you and need(s) something, then you have to find someone else to help you. [P21]

I would be finding one patient and then I would be distracted and go to another nurse who needed me. [P14]

The same HCP mentioned that the task is never just a “straight search” and that their presence on the floor would be noted by other patients and staff, and they would be drawn into monitoring or supporting interactions in support of patients or their colleagues.
Task-Technology Fit

Overview
This category of the FITT framework represents the working processes involved in completing the task (locating a patient) and how they interact with the RTLS (consisting of the bracelet, the wall-mounted beacons receiving the ultrawideband signal, the app on the tablet at the nursing desk, and the software application). The task-technology fit category consisted of three subthemes: (1) reorganization of the task, (2) personal control in relation to the task, and (3) efficiency and resource allocation.

Reorganization of the Task
Reorganization of the task was expressed by HCPs as the ability to locate and monitor patients in real time without having to go into each patient room. HCPs using the technology started their search at the nursing station, where they interacted with the tablet to identify the location of the patient and then proceeded directly to the patient, rather than conducting a random search. Providers found this beneficial for several reasons, including using the tablet to “spot-check” to see where a patient was or being able to directly find a patient when they wanted to:

The purpose as I see it, is to be able to locate the patient when you need to locate a patient, whether it’s because of double checking that they haven’t eloped or you’re trying to locate them for a therapeutic reason or for you know, I think this is the big one was like family members coming to visit and they wanted, you know, to find their loved one. [P4]

Furthermore, we found evidence of evolving work routines or stable actions that achieved work across time and space through adaptive routinization that supported HCPs use and adoption of the technology. However, adjustment of the technology (eg, tablet number and location) or routine (eg, shift change and anticipating interpatient alterations) to achieve a successful implementation was still clearly a work in progress at the time the interviews were conducted:

It’s been so new, it’s hard for us to remember that we have this...the more people are talking about it, the more that we’re remembering, “Oh we could use this instead of you know, running about the unit searching for patients.” [P3]

Habitually, I would start the lap- and then realize that I’m wasting time, cut through the care station, find them on the monitor and then, yeah. So, it-it-it did change the fact that I-I wouldn’t do more than one lap before realizing that I could go use it. [P8]

Right at shift change, I’m not gonna go in there and push through the nurses, trying to do their shift change to look at the monitor. I’ll just do a couple extra laps until I find them because it’s a reasonably small unit. [P8]

Personal Control in Relation to the Task
The second subtheme in the task-technology fit category was the recognition of personal control in relation to the task of locating patients. The HCPs perceived that they had more control by being able to decide how and when to find patients using the app. For example, if they were in the nursing station performing another task, they were also able to see or locate a patient at a glance on the tablet. Examples of this narrative are as follows:

Instead of walking around three times around the entire unit, I know exactly where my patient is. I also can monitor them, uh, remotely. [P5]

Especially, uh, at night-time it’s a lot better because, uh, we know that, uh, which patient is, uh, in the room and which one is out. When we are sitting in the nursing station, we can say like this...patient has come out of the room and is in a hallway or he has woken up. [P6]

At night, like I wanna see which patients are asleep, or awaken, are awake, like, if the bed alarms don’t work. [P13]

Efficiency and Resource Allocation
Efficiency and resource allocation were identified as a subtheme in the task-technology fit category. The HCPs stated that they were saving time and resources through their ability to find a patient directly using the app. They shared that a room-by-room search of the unit could potentially take up to 15 minutes each time as their patients were very mobile, which was challenging as they were also likely to be distracted by another patient during their search:

It saves me time from going from room to room because these patients, normally they walk around and go to other people’s rooms so where you least expect them to be that’s where they sometimes are. And, um, if you think they’re gonna be in their room, most times you’re wrong. They’re in somebody else’s room away from-- Maybe the other corner, you know, so it saves your time when you use this device. You located them faster. [P11]

HCPs further reflected that the time they saved using the app to locate patients could potentially be used for accomplishing other tasks. For example, 2 providers stated as follows:

When time is saved, then I can invest that time more on patient care, and all the things that I can do with extra time. [P13]

You still have to physically go look for the patient, but at least it does narrow it down for you. [P4]

However, the location of a singular tablet installed at the nursing station may pinpoint a patient in 1 location only to find that they had moved by the time the HCPs reached there. Finally, the technology was considered helpful in locating multiple patients when the providers needed to run a group activity:

If I run groups, and I want to get, you know, specific patients of a certain group—um, then I would, you know, go to iPad and see where they are. [P8]
Individual-Technology Fit

Overview

Individual-technology fit represents the relationship between the attributes of the individuals and the attributes of the technology and, in this study, comprised the most populous coding category for the interview data. The “individual” construct in the FITT model represents the HCP, who uses the technology to locate a patient; the patient, who wears the bracelet; the users as a group; and the organization, in which the technology is installed and the “tasks” and work processes are occurring. A total of four subthemes were identified in the individual-technology fit category: (1) privacy and personal agency, (2) trust in the technology, (3) user interfaces, and (4) perceptions of increased safety.

Privacy and Personal Agency

The subtheme of privacy and agency has multiple perspectives, given that while HCPs were the “users” of the technology, patients were the ones wearing the bracelet and being tracked. In alignment with the survey data, HCPs felt that patients on the unit had a limited right to privacy due to their perspective that these patients required close supervision to reduce their risk of harm (eg, preventing them from a fall or an altercation with another resident). This perception was expressed in narratives that suggested that patients with dementia lose their right to privacy (and conversely, their right to refuse to be monitored using the bracelet) as their cognitive impairment and behavior presents a risk to themselves, their family, and the organization:

*There is no privacy for patients with dementia, with behaviors...privacy means that you are putting them in, uh, at risk.* [P6]

However, HCPs recognized that other patients (eg, those without dementia) had a right to privacy and may thus find the use of RTLS to monitor them unacceptable:

*I don’t know about the patients on other floors if...it’s monitoring every move...they don’t wanna be like, to know that they’ve gone to the washroom, you know? These just—these are- these are things they might not like.* [P6]

Despite HCPs belief that patients in the unit waived their right to privacy in return for what they believed to be better care and safety, in general, providers believed they were “entitled to their own privacy” (P16) and did not agree with the concept of using location tracking for HCPs. In particular, HCPs were worried that if the ability to track staff was initiated by management, it could be used to sanction them:

*So, tracking staff, they would be depending on the culture, and I think just general trust of senior leadership. Like, are you tracking if I’m taking too many breaks, if I’m doing my job correctly?* [P21]

Those who personally felt that there may be value in tracking the movements of staff acknowledged that other HCPs might not agree but variously supported the tracking of providers to help locate each other, gather supporting data as it relates to workload monitoring, and build more efficient units. For example, they suggested that tracking providers’ movements may be useful for showing how much they move in a day, how often they are in patient rooms, how many staff members are required to assist a patient, and for how long:

*Cause I think we don’t estimate that correctly.* [P22]

The potential risks associated with the collection and storage of patients’ data were not a focus of concern; providers did not reflect on data use and storage until prompted during the interviews. Furthermore, providers made assumptions that the location data were kept on internal hospital servers; while for this study all data were stored on hospital secure servers, this is unusual and had to be negotiated with the equipment vendor who retained data on proprietary servers. One HCP stated as follows:

*I would hope the security, or the data is secure and, um, only, you know, used for the study purposes.* [P3]

Another wondered if the data could be used for alternative purposes that they were not aware of:

*I don’t know that information is being used for some other sinister reason I don’t know. Like whether they trust us with that information.* [P4]

Trust in the Technology

The subtheme of an individual’s trust in the technology relates to the reliability of the system—that the technology was working when required—and how it was used or adopted as a result. However, HCPs displayed a tolerance for unreliability in this product, likely because there were workarounds; that is, the providers could revert to walking the unit to find a patient, which made the consequences of unreliability minor:

*Okay. So, you have a little bit of a clue that there’s something wrong with the bracelet if it’s not moving then?* [Interviewer]

Yeah. Especially when the bracelet, uh, is on the screen in one location and patient is physically on a, in a completely different location. [P5]

Okay. Do you think that it not working, um, will influence how you trust the system or how you use it? [Interviewer]

*Uh, no, because it doesn’t happen often, it’s just once in a while malfunction.* [P5]

User Interfaces

The individual-technology fit category includes a subtheme of how patients and HCPs interfaced with the technology. Many HCPs shared that some patients “fidgeted” with or attempted to remove the bracelets. When asked to expand on this, HCPs explained that “some patients do remove, like they have catheter or other bracelets, they remove it. It’s just, uh, one of their behaviors” (P13). In response to a patient removing their RTLS bracelet or trying to remove it, providers reported that they would attempt to redirect the patient or move the bracelet to another location on their body (eg, an ankle) despite acknowledging that such attempts to remove the bracelet might be an expression of their desire to not wear the bracelet at all:
Sometimes we don’t understand, but of course they’re showing that they don’t like it. [P15]

Despite acknowledging this, HCPs did not explore the reasons why patients attempted to remove their bracelet. However, one provider did report that they would remove the bracelet if “the patient states exactly that he knows what the bracelet is and he states like, ‘I don’t want to be monitored”’ (P13).

Furthermore, providers mentioned that the patient-technology interface was a tool for maintaining patients’ dignity while wearing the bracelet, considering their lack of agency over whether they wore the tracking device. The bracelet esthetic mimicked something other than a surveillance device. For example, one provider reported that a patient “referred to the bracelet...as their Apple watch and it was kind of cool that they had a nice new watch, which was beautiful” (P07). She suggested that this was beneficial and stated as follows:

In the future, if it could even mimic a watch or have a watch face on it, that it, you know, when it goes into screensaver or whatever becomes a watch face, then it would be even more patient-friendly, user-friendly, right?” [P7]

Similarly, another HCP stated as follows:

Even people with dementia want to feel ready for the day and wear important pieces like a watch, carry a wallet. Wearing such pieces give dignity and acknowledges that people still have purpose. [P14]

Perceptions of Increased Safety

Another subtheme related to how providers supported and justified the use of the system by discussing potential (but not yet implemented) enhancements to the safety of patients. Specifically, providers felt that the technology might allow them to intervene or react more efficiently to a presumptive risk with the potential to avoid injury. One person in the study expressed an interest in using the RTLS to reduce the use of restraints, while others mentioned their desire for proximity alarms that would alert providers when patients who are “having an altercation” are in close proximity to each other and allow them to intervene. The narratives of this theme include the following:

In terms of safety, like interactions with other...other patients...like you could, you know, set up some kind of parameters that alert the staff when my patient X gets within vicinity of patient Y. [P20]

I think maybe even physically if you see that a patient has entered the washroom, an alert could go off into the care station. Um, there’s often times where I’ll find someone in there and they’re either beginning the process on their own and need help, or they are finishing the process and it’s become messy and there’s no way for us to know...But if an alert goes off in the care station, so and so has entered the bathroom, then whoever’s in the care station can respond to that before it gets messy. [P07]

Discussion

Principal Findings

In this study, the introduction of an RTLS with restricted functionality of tracking the location of people and assets within a secure inpatient unit for persons living with dementia represents one of the most elementary implementations of a single technology to achieve a single task. The successful adoption of the technology was uncomplicated, and its uptake by HCPs was thus predictably swift. We aimed to investigate the factors influencing the adoption or nonadoption of the RTLS technology. We found that the successful adoption of the RTLS was due to the strong fit between the technology and the task, which was locating the patients more efficiently, and the strong individual and task relationship on the unit.

First, we summarize the results of the interactions between the 3 pillars of the FITT framework to explain the success of the implementation and identify the factors that influenced the adoption of the technology. Second, we address related issues that warrant consideration for similar and more complex implementation.

Evaluating the FITT of the RTLS Implementation

The strong individual and task relationship on the unit was supported by an organization whose labor force was stressed due to the global pandemic and shortages and allied health and nursing HCPs who were aware of the challenges of ensuring that mobile individuals living with dementia were safe without constant one-to-one supervision. Individual providers who were responsible for locating patients in real time for events such as medication rounds, check-ins, patient visitors, and meals relied on installed closed-circuit television camera feeds. These cameras streamed to monitors mounted above the nursing station desk, providing real time but restricted line-of-sight hallway identification of patients and providers. However, the patient had to be in a hallway and remain in the same place long enough to be visualized. Depending on the patient, a care provider may need to search >1 bedroom to find a patient [22]. Providers’ anticipation of negative outcomes, both personally and to the patient from less than round-the-clock monitoring or to respond to risky behaviors such as interpersonal altercations or elopement, was a strong motivator not only to continue with what was described as an inefficient and time-consuming process but also to find and adopt a more efficient solution. As demonstrated in the narratives, the strong fit between the RTLS technology and the task was perceived as a clear benefit to staff and, as such, an influencing factor of technology uptake. When staff perceive that using the technology is the best fit for the completion of their tasks, they are more likely to adopt its use [23]. With this type of technology, Doshi-Velez et al [22] identified that being able to find patients more efficiently provides nursing staff with a strong motivation to use RTLS technology. Our HCPs provided examples of the perceived limitations of the current method of completing the patient locating. The time and energy required to find patients by “roaming” the halls were not generally perceived as a valuable use of time. Similar to research by Griffin et al [3], the HCPs appreciated the efficiencies that the technology afforded them.
As the RTLS technology in this study passively surveils, it is neither predictive nor prescriptive; therefore, while it may help identify potential issues (eg, a patient alone in a bathroom or exiting the facility), with the functionality activated for this pilot, the system still required a provider to identify an issue (by visualizing patient location on the tablet), judge the potential risk, and determine if they needed to respond in person. The transactional (locating task) and anticipatory (monitoring task) use of the technology were both highly suited to the task of reducing risk for the individual, the institution, and the patient and provided peace of mind to visiting care partners who may perceive that the providers were more aware of the presence of their family or loved one.

Another influencing factor for adoption was the fit between the individual and the technology, partly due to the intrinsic capacity of the HCPs. We conducted foundational preinstallation training on the operation and use of the RTLS, and the HCPs were comfortable with workflow changes as a result of being a research-intensive facility (eg, by way of social influence and facilitating conditions), which has been found to be beneficial for the adoption of technology [24,25]. In addition, as found in the preinterview survey, the HCPs self-reported a positive attitude toward technology in general and the RTLS in particular. However, the fact that the technology was introduced to the staff as a pilot research project may have impacted its adoption. One HCP noted that if it had been presented as the de facto and permanent new method of locating patients, they would have been more invested in it.

Similar to Griffin et al [3], the HCP’s confidence in the RTLS system was based on trust. This included trust that the organization would not implement a system that did not secure patient location data, that the RTLS was reliable and presented accurate information, and that the organization’s trust in them as professionals should preclude the expansion of RTLS technology to monitoring staff for any reason. Similarly, the trust extended to the reliability that the system would be functional when needed. While the implementation of the RTLS presented few technological issues, beyond unplugged beacons, frozen screens, and drained bracelet batteries, that were unable to be managed by the on-site research coordinator (in addition to its fit with the unit workflow and operations), other issues emerged that are noteworthy.

Perceptions of Safety

The decision to pilot or implement RTLS is frequently promoted as it has the potential to enhance unsupervised freedom of movement and improve the physical safety of persons living with dementia. Similar to the findings reported by Hall et al [26], the HCPs in our study identified that the primary rationale for using RTLS in this population was enhanced physical safety rather than freedom of movement. Providers alluded to technology-enhanced patient safety, yet it was difficult for them to articulate measurable outputs of related safety improvements compared to current interventions. It is important to distinguish between the “potential” of the technology’s capabilities and the actual functionality that helps realize improved safety. In reality, the functionality of the installed technology for our study was not “smart” nor was it predictive or responsive; unless a human was monitoring the app the moment a patient was at risk, there was no intervention or lessening of risk. The installed off-the-shelf technology in this study had both gofencing alerts and call button features but required human monitoring and sufficient staffing to respond and therefore were not enabled. Installations that enable these types of alerts (related to being outside a restricted area, experiencing a fall, or being in a location that suggests more risky behaviors) must be monitored to allow for a response in real time. This has resource implications related to technology (eg, mobile phone apps or tablets that allow remote monitoring rather than using a fixed desktop application) and human resources (eg, alerts must be responded to); thus, additional staff must be recruited to monitor, investigate, and respond to alerts anywhere on the floor at any time. Infrastructure and funding challenges and labor force shortages must be addressed to encourage more widespread exploration of the value and use of technologies such as RTLS in long-term care.

Training and Adoption

HCPs in the study were oriented to the tablet app through training sessions, which focused on the features of the technology and its basic use, whether they voluntarily chose to use the location map app or not. The tablet was located beside the legacy surveillance camera feed, serving as a reminder of its availability. During the training sessions, those who were using the technology shared use cases and examples of when and how they found the system useful. At the start of the project, HCPs explained that their confidence in the system would be heightened if they could trust that it reliably showed the real location of patients and that the patients were unable to interfere with its operation. By the end of the study, most HCPs were using the tablet app based on their perception of its usefulness, its alignment with their comfort level with technology, and its ability to help them perform job responsibilities. As the pilot implementation progressed, “superusers” emerged [27,28], who were early adopters and who strongly advocated for its usefulness and value in the unit. Superusers often spontaneously helped others troubleshoot or navigate the RTLS and quickly identified alternative uses for the technology (eg, as asset tags).

In one case, we downloaded a patient’s activity reports (a management feature of this particular RTLS), which helped to identify that they walked an extraordinary distance daily, allowing providers to integrate changes into the care plan (eg, encouraging rest times and increased caloric intake). Furthermore, some HCP users identified that the technology would be potentially more helpful if it were available on their
phones or if there were more tablets accessible throughout the unit at different locations.

The setting for this study was a locked inpatient unit that already used sensors fitted to each patient to prevent their unaccompanied exit from the unit. Within the unit, the patients were free to move as they wished or were able to. As a short-stay unit, its focus was on stabilizing or addressing behaviors and facilitating patients’ return to the community. Unlike a long-term care home where residents’ mobility is often compromised, many of the patients in this unit were ambulatory; the organization’s interest in preserving patients’ dignity and aligning their values with those of patients by allowing their free movement in the unit had operational implications, such as unsupervised interactions between patients and increased provider time in seeking out patients for treatment and well-being checks. It is unsurprising that organizations are moving to replace what they believe to be non–value-added tasks performed by providers with technology-enhanced solutions such as RTLS [29]. Patients’ resistance to wearing the bracelet was reduced for some patients perceived the technology similar to an Apple Watch, which they described as “cool.” Providers suggested that to avoid dissent, further disguising the transponder, what Sannon and Forte [30] describe as “dignity in design” where aesthetics are considered along with utility, would provide more dignity to their patients when the organization and substitute decision makers’ values (to create a safe and risk-free environment) supersede patients’ right to privacy.

During the training of HCPs on the RTLS and in our research interviews, we noted a lack of in-depth discussion regarding the complexities and potential challenges associated with implementing monitoring technologies in a health care setting [26]. This approach, while simplifying the training and implementation process, may have allowed staff to focus on its potential to enhance patient safety but failed to engage the HCPs, care partners, and, where feasible, patients in a dialogue about the implications and challenges associated with its deployment, such as its impact on privacy. Hall et al [26] suggested that for technology implementations to be successful, they must involve substantive discussions to anticipate and address these challenges. Such inclusive conversations involving all stakeholders in the decision-making process are crucial for a more effective implementation of novel technological systems.

In the survey, all participants identified that using this technology to track persons with dementia was either somewhat (16/20, 80%) or completely (4/20, 20%) acceptable. During the interviews, the staff did not initiate any ethical concerns. Prompted questions related to the ethical use of ubiquitous monitoring technologies that collect sensitive biometric data, the creation of a surveillance culture, and the responsibility of organizations to respect the rights and dignity of susceptible individuals when using these technologies were not identified as a care priority by staff in the interviews. This finding has considerable risks and implications that have been explored in more detail elsewhere [2].

In the survey, the HCPs demonstrated a mixed level of acceptance of the RTLS being used to monitor HCP movement. This finding was echoed in the interviews, where HCPs expressed a hesitancy to use the technologies on themselves despite overall satisfaction with their work; this is consistent with previous studies [3,31], where workforce monitoring was viewed as a lack of management-worker trust, a manifestation of the blame culture, and a foot in the door where monitoring for quality of care (eg, time spent with patients or handwashing) would be the first step to its use for individual performance tracking (eg, length of breaks and productivity). The shift of health care administration, toward the adoption of more scientific management where standardization, removal of inefficiencies, and process improvement, supports the notion that RTLS might be used in this fashion [32]. Overt messaging from the management that the monitoring in this study was limited to patients supported its successful adoption and implementation.

**HCP Workload and Technology Support**

Technology has increased the amount and complexity of information that employees are expected to process and has enabled HCPs to access unlimited amounts of information to do their jobs [33]. In the health care environment in which provider resources are scarce, workloads are high. Some HCPs on the floor did not use the RTLS (and did not volunteer to participate in our study); those who participated cited its simplicity, reliability, and ability to reduce their perceived workload as the reason for their continued use of the technology, rather than the rationale of safety as the rationale for adoption. Workload presented a challenge in relation to troubleshooting technological issues with the RTLS. Although providers were comfortable with the RTLS, none moved into a role where they were able to troubleshoot simple technological issues. Most cited the lack of time, the availability of a low-cost alternative (walking around), and the presence of a research team for technological support; however, this support was not available 24/7.

**Limitations**

The generalization of findings from data collected at a single location during one technology implementation has limitations. However, our focus on exploring the perspectives of HCPs working in an environment that provides specialized dementia care was both a strength and a limitation. The perspectives of the organizational decision makers on the rationale for adopting and implementing the RTLS technology in the unit were outside the scope of the study. These findings will be important to explore in future research on the adoption of RTLS in dementia care settings.

Providers’ self-reported comfort with the installed RTLS was evident in the preinterview questionnaire responses and after providers were trained and had used the technology. Furthermore, the unit under study is a teaching and research-intensive hospital that frequently involves providers in technology-enabled research studies. The uniqueness of the unit may have impacted the self-reported providers’ job satisfaction where, particularly during the pandemic, restricted access to the unit and the limited number of patients and providers offered some immunity from the ongoing pandemic pressures related to staffing shortages, frequent outbreaks, and
increased workloads, which were experienced in other congregate settings for people living with dementia [34].

The staff members on the unit were acquainted with the principal investigator and the research coordinator of this study, who conveyed the study’s purpose to the participants, responded to staff inquiries, and obtained participant consent. To ensure impartiality, the study team used deductive analysis of the data and applied the FITT framework as a standardized approach to analyze the data. This framework is a well-established and widely used approach in the literature, which helped to reduce the potential for researcher bias in data interpretation. Furthermore, the fact that the unit uses other technology, such as the elopement prevention system, suggests that staff members are accustomed to using technology in their daily work routine. This familiarity with technology may have helped to reduce any bias toward or against the RTLS technology used in this study.

Another limitation relates to the small scale of this study. Although the number of participants was sufficient to reach saturation with respect to our research questions, it is insufficient to generalize beyond the specific context of the featured location-tracking technology, its functionality, and the hospital-based dementia care setting in which it was implemented.

Conclusions

Similar to many novel technologies in the nascent stages of their adoption, evidence to support the utility and effectiveness of RTLS in improving the safety and quality of care in health care organizations and patients’ experience is limited. While HCPs were able to reduce their self-reported patient search time dramatically, sometimes by half, any improvements in care were implied or perceived. While no participants self-reported evidence of reduced patient risk, reduced adverse events, or improved outcomes, they described greater perceived peace of mind for the staff responsible for oversight. As stewards of resource-constrained pragmatic organizations, decision makers in the health care sector will weigh the risks of RTLS adoption related to personal privacy, overreliance on untested technology, and cost against the benefits of ubiquitous monitoring of human and equipment assets, performance management, and automation of location tasks to improve staff efficiency. The decision to adopt novel technologies necessitates examination policies, workflows, and resource commitments beyond the initial purchase costs of the hardware, software, and training to identify technologies and adoption processes that best fit the organizational context and the tasks it must perform.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide.

[DOCX File, 38 KB - aging_v7i1e45978_app1.docx ]

References


Abbreviations

- FITT: Fit Between Individuals, Tasks, and Technology
- HCP: health care provider
- REDCap: Research Electronic Data Capture
- RTLS: real-time location system
- UHN: University Health Network

Edited by R Yang; submitted 24.01.23; peer-reviewed by M Molinari Ulate, MM Qirtas; comments to author 20.03.23; revised version received 15.05.23; accepted 27.02.24; published 08.04.24.

Please cite as:
Factors That Influence Successful Adoption of Real-Time Location Systems for Use in a Dementia Care Setting: Mixed Methods Study
JMIR Aging 2024;7:e45978
URL: https://aging.jmir.org/2024/1/e45978
doi:10.2196/45978
PMID:38587884

©Lynn Haslam-Larmer, Alisa Grigorovich, Leia Shum, Andria Bianchi, Kristine Newman, Andrea Iaboni, Josephine McMurray. Originally published in JMIR Aging (https://aging.jmir.org), 08.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Using the TrueLoo Smart Device to Record Toileting Sessions in Older Adults: Retrospective Validation and Acceptance Study

Jordan Glenn¹, PhD; Parmoon Sarmadi², MEng; Paul Cristman², PhD; Gabrielle Kim², BASc; Ting-Hsuan Lin², MA; Vikram Kashyap², AB

¹Department of Health, Human Performance and Recreation, University of Arkansas, Fayetteville, AR, United States
²Toi Labs, Inc, San Francisco, CA, United States

Corresponding Author:
Jordan Glenn, PhD
Department of Health, Human Performance and Recreation
University of Arkansas
155 Stadium Dr
Fayetteville, AR, 72701
United States
Phone: 1 561 513 3865
Email: jordan.mckenzie.glenn@gmail.com

Abstract

Background: Because of the relationship between independent living and activities of daily living, care teams spend significant time managing assisted living residents’ toileting problems. Recently, the TrueLoo was developed as a connected toilet seat to automatically log and monitor toileting sessions.

Objective: This study aimed to demonstrate the validity of the TrueLoo to (1) record and identify toileting sessions with regard to stool and urine events; (2) compare the results with the person-reported, standard-of-care methods; and (3) establish metrics of user acceptability and ease of use in a assisted living facility population.

Methods: We used two phases: (1) initial development of the TrueLoo algorithms to accurately identify urine and stool events and (2) evaluation of the algorithms against person-reported, standard-of-care methods commonly used in assisted living facilities. Phase 2 analyzed data over a 3-day period from 52 devices. Participants’ age ranged from 63 to 101 (mean 84, SD 9.35) years. Acceptability and ease-of-use data were also collected.

Results: Regarding the development of the TrueLoo algorithm for urine assessment, sensitivity and specificity of 96% and 85% were observed when evaluating a gold-standard labeled data set, respectively ($F_1$-score=0.95). For stool, sensitivity and specificity of 90% and 79% were observed, respectively ($F_1$-score=0.85). Regarding the TrueLoo algorithm in assisted living settings, classification performance statistics for urine assessment revealed sensitivity and specificity of 84% and 94%, respectively ($F_1$-score=0.90), and for stool, 92% and 98%, respectively ($F_1$-score=0.91). Throughout the study, 46 person-reported instances of urine were documented, compared with 630 recorded by the TrueLoo. For stool events, 116 person-reported events were reported, compared with 153 by the TrueLoo. This indicates that person-reported events were captured 7% (46/630) of the time for urine and 76% (116/153) of the time for stool. Overall, 45% (32/71) of participants said that the new toilet seat was better than their previous one, 84% (60/71) reported that using the TrueLoo was easy, and 99% (69/71) said that they believed the system could help aging adults. Over 98% (69/71) of participants reported that they would find alerts related to their health valuable and would be willing to share this information with their doctor. When asked about sharing information with caregivers, 66% (46/71) reported that they would prefer the TrueLoo to send information and alerts to their caregiver, as opposed to the participant having to personally communicate those details.

Conclusions: The TrueLoo accurately recorded toileting sessions compared with standard-of-care methods, successfully establishing metrics of user acceptability and ease of use in assisted living populations. While additional validation studies are warranted, data presented in this paper support the use of the TrueLoo in assisted living settings as a model of event monitoring during toileting.

(JMIR Aging 2024;7:e50856) doi:10.2196/50856
KEYWORDS
activities of daily living; toileting habits; bowel movements; elder care; smart toileting; monitoring technology

Introduction

The global population is aging rapidly, and by the year 2050, the number of people aged 60 years and older will drastically increase to over 2 billion [1]. Concurrent with this increase in the number of older individuals, life expectancy is projected to increase in the coming decades [2]. However, as is often the case with advanced aging, this increased life expectancy will almost certainly be associated with accompanying morbidities and their associated costs to individuals and health systems [3,4].

These aging demographics will increase the burden of care in assisted living facilities. As part of their regular responsibilities, care staff in these facilities spend significant time and effort documenting and managing their residents’ toileting habits and events [5,6]. This also means they play a key role in the assessment and prevention of toileting-related issues.

While toileting measurement and assistance is an important part of management in assisted living facilities, there is little evidence on how these individuals actually manage bowel problems and the quality of the interventions performed [7]. Furthermore, to track these events and monitor for issues, staff often rely on human reporting. This can come in the form of second-hand information from the resident or the staff member’s own memory when recording multiple residents’ toileting habits at once near the end of a shift; as a result, these methods are incomplete and prone to error.

Because of the relationship between independent living and activities of daily living [8], care teams spend a significant amount of time managing residents’ bowel and bladder problems. Besides common bladder issues such as urinary tract infections [9], bowel problems such as constipation [10], diarrhea [11], and fecal incontinence [12] are highly prevalent among assisted living residents and present several challenges for care staff with regard to monitoring and reporting. Furthermore, these types of events cannot be identified without proper monitoring and recording of toileting events. Recently, the TrueLoo (TL; Toi Labs) was developed as a connected toilet seat to automatically log and monitor toileting sessions, removing the burden from the patient or facility staff. Reporting and monitoring stool and urine characteristics is shown to notably improve the quality of care that residents in assisted living facilities receive, especially those living with multiple comorbidities [13]. The TrueLoo provides concurrent monitoring of toileting sessions, providing caregivers with data to improve their clinical decision-making through evidence-based technology.

Previous investigations and review papers have evaluated the efficacy, practicality, and use of smart toilet seats as a model to longitudinally monitor individual toileting habits. Initial work provided evidence of usability and proof of concept for such a design through the use of a colorimetric assay tracing red-green-blue values from images of urinalysis strips [14]. They also included cameras to collect “analprints” used as unique identifiers. However, there were limitations related to device scalability as well as privacy from upward facing cameras; nevertheless, this was important work to demonstrate functionality of the process [14]. Further work demonstrated the ability to use an image-based data set to classify stool according to the well-established Bristol Stool Form Scale [15], demonstrating the feasibility of using such a technique to automatically evaluate samples [16]. However, while the groundwork for using smart toileting technology has been demonstrated, no data currently exist demonstrating the efficacy and applicability of accurately capturing and evaluating toileting events in a real-world setting.

Therefore, the purpose of this retrospective study was to demonstrate the validity of the TrueLoo (1) record and identify toileting sessions with regard to stool and urine events; (2) compare the results with the person-reported, standard-of-care methods; and (3) establish metrics of user acceptability and ease of use in an assisted living facility population. We hypothesized that the TrueLoo would demonstrate significant improvements in the event capture of toileting sessions compared with the current reporting systems, while being well accepted in the target population.

Methods

Ethical Considerations

This was a retrospective analysis, and all procedures were approved by the Western Institutional Review Board (TLSD-001); the study was conducted in accordance with the Declaration of Helsinki.

Subject Demographics and Recruitment

The participants in the study were aged between 63 and 101 (mean 84, SD 9.35) years. All study participants were from a skilled nursing facility located in Dallas, Texas. Of the 52 participants, 27 (52%) were female. The eligible participants were contacted by community staff, and all individuals who agreed to participate in this investigation were provided with an extensive overview of the product and its capabilities; there were no residents who declined participation. The use of the technology was noted in each resident’s care plan, and community staff were extensively trained by Toi Labs on the capabilities of the product. Finally, Toi Labs provided additional materials to the community to distribute to residents and their families to address concerns about privacy protocols, product overview questions, or other questions regarding the reporting system.

Inclusion and Exclusion Criteria

This study was conducted as a retrospective analysis, and, as such, there was no participant consent required. However, to participate in the original commercial program, each participant was required to meet all of the following inclusion criteria: (1) willing to participate and provide consent for the program, (2) aged 55 years or older, (3) a resident of the assisted living
facility where the data were collected, and (4) had regular access to a TrueLoo.

Additionally, participants satisfying any of the following exclusion criteria were precluded from participation in the study: (1) unwilling or unable to accept the requirements associated with installing the TrueLoo in their residence, including power and Wi-Fi connectivity, and (2) used certain types of toileting assistance devices that, at this time, are not compatible with the TrueLoo (e.g., padded toilet seat risers).

Study Design
This study was conducted in two phases:

1. Development of the TrueLoo algorithms to accurately identify urine and stool events
2. Evaluation of the algorithms against the real-world person-reported, standard-of-care methods commonly used in assisted living facilities

It is suggested that traditional methods for developing robust evidence are incongruent with the agile approach commonly used in software development, as there is often an incongruence between the length of traditional registered controlled trials and the development and update cycle for software [17]. Given the agreement that more traditional approaches present fundamental limitations for proper evidence generation for digital health solutions, this 2-phase approach allowed for the controlled development of the algorithms from real-world data collected as part of a commercial program and then analyzed and evaluated retrospectively for efficacy.

For phase 1, details surrounding the development of the initial algorithm for identifying stool and urine events are described below. For phase 2 (real-world data collection), participants were enrolled independent of symptoms, disease diagnosis, or state. Furthermore, in an effort to attain real-world applicability, there was no stratification of the population by sex, age, race, or disease severity. Each participant was monitored in their residences within their respective community. For each resident, a TrueLoo toilet seat was installed in their private bathroom along with an initial health assessment to collect demographic information and any preexisting conditions. During this phase, 52 devices were deployed and reviewed for a 3-day period. Data were collected on each of the 52 participants using the TrueLoo.

Measures

Development of the TrueLoo Algorithm for Identifying Urine and Stool Events
To verify urine and stool events that were accurately captured in the development of the algorithm, a panel of board-certified gastroenterologist, subject-matter experts were first enlisted to create a “gold-standard” database. The subject-matter experts created a rule set for image annotators, who were trained to accurately identify image content. These image annotators provided our ground truth and produced a labeled data set that was then used to train the machine learning algorithms before being run on the full data set. The applied labels were used to create digital biomarkers for toileting event imagery. As of this investigation, more than 10,000 sessions (times people have used the toilet) were labeled, with more than 40,000 images. A session was activated when a user is in proximity to or sitting on the toilet seat. Each session comprises multiple images that depict the physical characteristics of urine and stool seen throughout the session in a time series.

The TrueLoo algorithms were developed using an existing large image model that is pretrained to extract fundamental image features such as shapes, colors, and textures. This model is then trained on proprietary image data collected by the TrueLoo that have been human annotated with label taxonomy of over 20 contextual labels that include relevant health metrics such as the Bristol Stool Scale and nonhealth metric such as toilet cleaning. The labeled data set is used to refine the large image model to create the proprietary TrueLoo algorithms.

A classic deep learning network structure was used for the neural network architecture, as illustrated in Figure 1. The network consists of 5 convolutional blocks with max pooling for feature learning and extraction, followed by 2 dense layers. A sigmoid activation function on the final layer is used for the final multilabel classification task and a binary cross-entropy loss function. This architecture has been very successful in traditional image classification tasks, and pretrained weight configurations based on classic benchmarks are readily available, making it an ideal candidate for fine-tuning and transfer learning.
Detailed Working Description of the TrueLoo Technology

The TrueLoo consists of 2 parts: a hardware component that is delivered as a replacement toilet seat and a software system for analysis and reporting. Figure 2A shows a photograph of the TrueLoo seat with features called out and an image (example in Figure 2B) captured from the optical system. The TrueLoo seat has 2 user presence sensors: a contact sensor bound to the seat with no visible sign to the user and a noncontact time-of-flight distance sensor (Figure 2A.1) that activates when the user does not sit on the seat (ie, standing while urinating). By using the 2 sensors, the system distinguishes between standing and seated events as well as nonevent classification. The rear housing is used to mount the optical system and support electronics. The bowl is illuminated (Figure 2A.3) uniformly by red-green-blue-white LEDs to control color balance and some narrow band imaging illuminating with only 1 color. This allows for consistent imaging conditions for all currently encountered toilet geometries. Not shown in the image is the red-green-blue 8-megapixel manual focused camera and needed control and communication electronics. The system is powered by a single-board computer with integrated Wi-Fi communications for transmitting the images. The TrueLoo seat has a guest button (Figure 2A.2) to disable the system if a guest needs to use the toilet. The guest button automatically resets after each use. No images are recorded from guest events; however, they are registered in the database as an activation of the TrueLoo. Not shown in the figure is the cable routing using a conduit to fix the cable to the wall and connect the unit to power using a wall mount type AC-DC transformer at the outlet allowing for a long cable run with a low-voltage thin wire; this setup does not require a new outlet or the replacement or recharging of batteries. The seat is fixed to the toilet using a standard commercial mounting system for replacement toilet seats. After the TrueLoo seat is installed, it requires minimal to no ongoing maintenance, other than ensuring that the optics stay clean.
When a user activates the TrueLoo by sitting on the seat or standing in front of the toilet, the system activates an event and immediately starts imaging at 1.2 frames per second. The TrueLoo continuously captures images of the bowl for the duration of time the user is seated or standing in front of the device. Immediately after the event is finished the images are transferred via Wi-Fi to HIPAA-compliant servers for storage and analysis (Figure 2C). Duplicate images are not uploaded but are registered to ensure correct time sequencing and more.

Because of the non–battery power and Wi-Fi connectivity configuration, data could be logged and monitored to ensure that the device was connected and working properly. The device can save approximately 1 week of data locally and resume uploading data if and when a Wi-Fi connection is compromised. This capability provides added capture and integrity in the event of a prolonged Wi-Fi outage.

**Tracking of Toileting Events in Assisted Living Facilities**

As previously mentioned, person-reported methods of bowel movement and urinary event tracking are considered the standard of care in assisted living settings. For phase 2 of this study, the same annotators retrospectively labeled toileting images captured through a commercial engagement, creating a real-world ground truth data set. We then analyzed recorded TL-captured events through the algorithm developed in phase 1, comparing them with the person-reported events captured by facility staff, broken down into urinations and bowel movements.

The current practice for reporting toileting events among facility staff involves documentation at the end of a shift, sometimes manually (pen to paper in a chart), or other times in the electronic medical record. The analysis done was for a community that reports this information in the electronic medical record. The data recorded in bowel and bladder logs are often inadequate, with limited description. For example, the logging for urination only allows facility staff to notate “Void? Yes or No.” Many of the entries are listed as “Not Applicable.”

Bowel elimination questions include (1) the size of bowel movement (small, medium, large, resident not available, resident refused, and not applicable), (2) consistency of bowel movement (formed or normal, loose or diarrhea, constipation or hard, putty like, resident not available, resident refused, and not applicable), and (3) bowel continence (continent, incontinent, no bowel movement, continence not rated due to ostomy, resident not available, resident refused, and not applicable).

The TrueLoo system automatically classifies on an established scale based on the Bristol Stool Scale [15]. The following classifications were used to inform the TrueLoo algorithms: separate, hard lump nut-like stool; soft blobs with distinct edges; sausage-like stool with surface cracks; lumpy, sausage-like stool; smooth, sausage-snake stool; fluffy, mushy stool with ragged edges; and watery liquid, no solid stool.

Most of the data captured by the TrueLoo are supplemental when benchmarked against the information recorded by facility staff through the logs. As seen in our analysis, although the elimination questions in the electronic records have the potential to capture the characteristics of stool and urine, they are not accurately completed by care staff, if at all.

**Data Safety and Integrity**

Protocols are included as a part of the TrueLoo implementation’s standard process to ensure that privacy is maintained at all times through the data collection process.

1. The TrueLoo imaging system faces down into the toilet bowl and is designed specifically to scan stool and urine; it does not capture any body parts.
2. User information is completely deidentified, from data capture to analysis.
3. The seat itself does not carry the name or location of the user.
4. All deidentified data are stored internally—none of the data that are viewed or used are associated with any sort of identifier—and are completely anonymized.

Secure servers and connections allow reports to only be shared with onsite care teams tasked with caring for the user. The reports can be shared with outside care partners (such as physicians) or families. Toi Labs’ team members review and
validate reports to ensure compliance of privacy protocols and offer the best service and care to users.

Information is not identifiable on capture. Multiple protocols are in place throughout data capture to ensure that the data remain completely deidentified. No TrueLoo seat or its data are directly linked to a resident’s name or room number. This information is secured internally and is only used when providing reports to care teams. Otherwise, all data that are captured and uploaded are deidentified and anonymized. Secure servers and connections allow reports to only be shared with onsite care teams tasked with caring for the user. Upon the authorization of the user or their responsible party, the reports can be shared with outside care partners or families as they see fit. The reports are not distributed to anyone outside of the circle of care of the user.

Participant Satisfaction and Acceptability
At the end of the study, a survey about user satisfaction was provided to all participants to evaluate acceptability and comfort with the use of the TrueLoo. This survey can be found in Multimedia Appendix 1.

Data Analyses
Python version 3.9 (Python Software Foundation) and R version 4.1.3 (R Foundation) were used to conduct all analyses. Descriptive statistics were calculated for age and sex. For the development of the initial algorithms, sensitivity, specificity, precision, recall, $F_1$-scores, and receiver operating characteristic (ROC) area-under-the-curve (AUC) were calculated to evaluate the classification performance of stool and urine events. Bootstrapping was used to calculate 95% confidence intervals for AUC.

For evaluation of the algorithm in real-world evidence (RWE) settings, sensitivity, specificity, precision, recall, $F_1$-scores, and ROC analyses were performed. Overall results were compared with the person-reported events as a percent to determine effectiveness to the standard of care.

Results

Data Characteristics
Throughout the period involved in the study’s retrospective analysis, a total of 645 toileting sessions were recorded by the TrueLoo. Of the 645 total recorded events, 630 included urine and 153 included stool. There was overlap between the number of sessions containing both urine and stool. In this investigation, the average age of the participants was 84 (SD 9) years, with 52% (27/52) of the participants identifying as female. All of the individuals in the setting in question used the TrueLoo for the 3 days as planned. No participants declined the use of the TrueLoo.

Development of the TrueLoo Algorithm for Identifying Urine and Stool Events
With regard to the development of the TrueLoo algorithm, classification performance statistics for all urine and stool events can be found in Table 1, and ROC curves can be found in Figure 3. For urine assessment, AUC was 0.92, with sensitivity and specificity of 96% and 85% observed, respectively. For stool, AUC was 0.96, with sensitivity and specificity of 90% and 79% observed, respectively.

Table 1. Classification performance of the TrueLoo algorithm for detecting urine and stool events.

<table>
<thead>
<tr>
<th>Event</th>
<th>Precision</th>
<th>Recall</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>$F_1$-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urine</td>
<td>0.94</td>
<td>0.96</td>
<td>0.96</td>
<td>0.85</td>
<td>0.95</td>
</tr>
<tr>
<td>Stool</td>
<td>0.80</td>
<td>0.90</td>
<td>0.90</td>
<td>0.79</td>
<td>0.85</td>
</tr>
</tbody>
</table>

Figure 3. Receiver operating characteristic area-under-the-curve (AUC) analysis for the TrueLoo algorithm to identify (A) urine and (B) stool events. fpr: false positive rate; tpr: true positive rate.
Toileting Sessions Compared With Standard of Care

With regard to the evaluation of the TrueLoo algorithm in RWE settings, classification performance statistics for all events can be found in Table 2, while ROC curves can be found in Figure 4. For urine assessment, AUC was 0.95, with sensitivity and specificity of 84% and 94% observed, respectively. For stool, AUC was 0.98, with sensitivity and specificity of 92% and 98% observed, respectively.

Table 2. Classification performance of TrueLoo for detecting urine and stool events in assisted living settings (real-world evidence analysis).

<table>
<thead>
<tr>
<th>Event</th>
<th>Precision</th>
<th>Recall</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>F1-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urine</td>
<td>0.97</td>
<td>0.84</td>
<td>0.84</td>
<td>0.94</td>
<td>0.90</td>
</tr>
<tr>
<td>Stool</td>
<td>0.91</td>
<td>0.92</td>
<td>0.92</td>
<td>0.98</td>
<td>0.91</td>
</tr>
</tbody>
</table>

Figure 4. Receiver operating characteristic area-under-the-curve (AUC) analysis for the TrueLoo algorithm in real-world evidence settings (ie, standard of care) to identify (A) urine and (B) stool events. fpr: false positive rate; tpr: true positive rate.

To compare the TrueLoo with the person-reported standard of care, we compared the number of reported instances in the skilled nursing facility with the number of events recorded by the TrueLoo. Over the course of the study, there were 46 person-reported instances of urine documented, compared with 630 by the TrueLoo. For stool events, there were 116 person-reported events compared with 153 events reported by the TrueLoo. This indicates that, when compared with the TL, person-reported events were captured 7% (46/630) of the time for urine and 76% (116/153) of the time for stool.

Survey Feedback on Acceptability and Ease of Use

A high degree of user satisfaction was found in the exit survey. Overall, 45% (32/71) of participants said the new toilet seat was better than their previous seat, while 46% (33/71) said it was the same. Importantly, 84% (60/71) of participants reported that using the TrueLoo was easy and required no effort; 99% (69/71) said they believed that the monitoring system had the potential to help aging adults.

With regard to the transmitting and interpretation of data, 99% (69/71) of participants reported that they would find alerts related to their health valuable and would also be willing to share this information with their doctor. When asked about sharing information with caregivers, 66% (46/71) reported that they would prefer the TrueLoo to send information and alerts to their caregiver, as opposed to the participant having to personally communicate those details.

Discussion

Principal Findings

The purpose of this retrospective study was to demonstrate the validity of the TrueLoo to (1) record and identify toileting sessions with regard to stool and urine events; (2) compare the results with the person-reported, standard-of-care methods; and (3) establish metrics of user acceptability and ease of use in a assisted living facility population. Our hypothesis that the TrueLoo would accurately identify urine and stool events, as well as capture more toileting sessions than were captured via the standard of care, was confirmed. We also confirmed our hypothesis that this method would be well accepted by residents.

Ability to Record and Identify Toileting Sessions

In this investigation, the TrueLoo algorithm demonstrated high sensitivity, specificity, and accuracy for detecting urine and stool events in real-world settings. Toileting issues such as urinary tract infections [9], constipation [10], diarrhea [11], and fecal incontinence [12] are highly prevalent among assisted living residents; however, if regular monitoring of toileting events are not accurate, it becomes difficult to identify and manage these conditions. Being able to manage these clinical events begins with the identification and monitoring of toileting; if it cannot be tracked, it cannot be measured and therefore cannot be improved.

Previous research investigating the use of smart toileting technology provides an initial look into the feasibility and proof of concept related to successfully executing this type of
monitoring model [14-16]. However, these designs and models were limited as it pertains to real-world applicability. Early work used a colorimetric assay tracing red-green-blue values from images of urinalysis strips as well as upward facing cameras to collect “analprints” used as unique identifiers [14]. Based on the recommendation by Ge et al [18], “To enhance data quality, devices should be designed in ways that are physically or psychologically unobtrusive so as not to influence normal toileting behavior,” this initial model becomes problematic as the process can be intrusive and pose risks to data privacy.

Comparison of TrueLoo With Traditional Standard of Care

The standard of care (facility staff manual reporting) requires an individual to remember each event, manually track the event, and recall specific details about the event that would be relevant to a clinical issue. While this is theoretically feasible, it is unrealistic to expect these individuals to remember specific details and document each event without error throughout the course of a shift. Furthermore, expecting patients to remember and accurately report their own events is difficult due to recall bias and the natural discomfort of discussing one’s own toileting habits [19], especially if the events are different or unusual. This can be combined with the fact that, perhaps unsurprisingly, residents in seniors living facilities report that noninvasive methods of care are preferable to more invasive methods when it comes to the development of toileting programs [19,20]. This concept that there is a reluctance to openly discuss excreta [21,22] has hampered the development and acceptance of smart toilets; however, the use of smart toileting technology, such as the TL, creates a way to seamlessly integrate toileting analyses as part of routine monitoring, serving as a gateway to the digitalization of health care in the home [18].

Through this investigation, we found that facility staff underreport toileting events when compared with the TrueLoo. Given the active versus passive nature of the 2 methods, this is unremarkable, provided the limitations naturally inherent on the time and resources of human monitoring. Not only are facility staff dependent on their own timing for successfully monitoring a resident’s toileting habits, but they are also highly reliant on residents being honest about their own habits in the instances when they cannot be monitored or observed directly [13]. Furthermore, it appears that staff tend to log these data points in batches, often from memory. For example, most sites use a shift system where their staff are working from 6:00 to 14:00 hours, from 14:00 to 23:00 hours, and from 23:00 to 6:00 hours. Based on a single day of data provided by the site from their self-reporting logs (Multimedia Appendix 2), the majority of toileting events are reported toward the end of their respective shift. This is especially evident in the later evening shifts where the majority of reports from the 14:00-23:00 shift were reported from memory within a 2-hour time span. The level of cognitive load required with this method, in addition to their other responsibilities [7], is highly prone to error.

Many assisted living residents require more complex and advanced care compared with the general population, and evidence supports the fact that nursing home clinical outcomes are heavily reliant on geriatric approaches and care leadership. Proper care of these residents requires a multidimensional and specialized approach from facility staff [23,24]; however, recognition for this type of skill and the effort involved is frequently undervalued, thus leading to staff turnover and vacancies [25]. Furthermore, the high prevalence of toileting issues among nursing home residents indicates that there is demonstrable potential for improvement within this population. Such issues among residents are largely related to remediable factors, which can potentially be prevented or improved [7], and current nursing home practices do not adequately address these challenges. Providing a passive monitoring toileting solution that captures these types of data automatically, analyzes them, and transmits them back to the facility can reduce time and discomfort required for staff. This would allow an already underappreciated group [25] to focus on alternative needs in their respective facilities, removing burden and potentially increasing staff morale and attitude. The latter is, perhaps, most important because previous research from assisted living settings has shown that attitudes of care staff toward their organization, residents, and families have a significant effect on the quality of care provided to the residents [26].

Clinical Relevance in Real-World Settings

As discussed above, current standard-of-care methods are error prone and inconsistent. This inconsistency of monitoring potentially creates a larger issue for identifying critical conditions associated with greater costs and health care needs. Residents in assisted living facilities are at significant risk of developing issues such as urinary tract infections [9] and bowel problems such as constipation [10], diarrhea [11], and fecal incontinence [12], to name a few. These issues cannot be consistently identified without proper monitoring of toileting events. For example, when evaluating the prevalence of inpatient falls in a Michigan community hospital, 45.2% were related to toileting-based issues [27]. Importantly, 82.3% of patients who fell had completed a fall risk assessment before the incidents, indicating minimal relationship between fall assessment and actual falls. These results were further confirmed in a secondary data analysis conducted on 281,865 high-risk falls assessments collected in a multisite study where toileting issues were the third most powerful predictor of falls after “falls in the last 6 months” and “confusion” [28]. Furthermore, a retrospective analysis of falls related to nighttime toileting over a 1-year period found that 34% of falls were related to toileting-related issues [29]. Finally, a cross-sectional report evaluating the association between toileting and falls in older adults admitted to an emergency department discovered that the rate of recurrent falls was significantly higher in a toileting-related falls group than a non-toileting-related falls group [30]. As such, it is critical that toileting issues be addressed upstream to prevent larger, related issues from occurring. The use of the TrueLoo to accurately, consistently, and passively track toileting habits and related issues may provide the necessary feedback to ensure that residents receive the care and attention required to prevent falls from occurring, instead of requiring active response after the fall occurs.
Acceptability and Ease of Use

There are an ever-growing number of technological solutions offering potential benefits for older adults. However, despite the potential benefits, older adults regularly demonstrate lower adoption rates compared with their younger peers [31-33]. In the older adult population, perceived value, confidence in the ability to learn the technology, and the perceived impact on quality of life are reported to be some of the most robust predictors of willingness to adopt technology [34]. In this investigation, 84% (60/71) participants reported that using the TrueLoo was easy and required no effort (ability to learn) and 99% (69/71) said they believed that the monitoring system had the potential to help aging adults (perceived impact). Additionally, 99% (69/71) of participants reported that they would find alerts related to their health valuable and would also be willing to share this information with their doctor, and 66% (46/71) reported that they would prefer the TrueLoo to send information and alerts to their caregiver, as opposed to the participant having to personally communicate those details (quality-of-life improvements). These data points triangulate to the TrueLoo being not only efficacious but also successfully adopted as a passive monitoring intervention in this age group.

Strengths and Limitations

The strengths of this study include the process of initially developing the algorithm for identifying stool and urine events against a gold-standard labeled data set and then retrospectively analyzing it in an RWE setting. This methodology allowed for the ability to generate real-world insights into how the TrueLoo could most effectively be used in assisted living facilities. Additionally, the ability to get direct feedback from the end users provides subjective validation, in addition to the efficacy of the TrueLoo to evaluate sessions. There is often a disconnect between clinical efficacy and practical use, which prevents new technology from being properly implemented. While RWE designs are powerful for real-world practicality, there are associated limitations. In this study, there was a lack of ability to control certain parameters (monitoring and reporting habits of facility staff) and collect certain data points (detailed health reports) on the participants. Ameliorating these limitations would change facility workflows, therefore affecting the validity of real-world efficacy. As such, we accept these limitations but recommend that this study be followed up with additional controlled investigations into the clinical efficacy of the TrueLoo. Such controlled investigations would also allow for the evaluation of other indications such as loose or bloody stool or cloudy urine. Given the clinical applicability of these indications, they are a recommended next step for future research.

Conclusions

In this retrospective validation and acceptance study, we demonstrated the validity of the TrueLoo to record toileting sessions compared with the standard-of-care methods, while categorizing them into clinically relevant events. Additionally, the TrueLoo successfully established metrics of user acceptability and ease of use in assisted living populations. While additional validation studies are warranted, the data presented in this paper support the use of the TrueLoo in assisted living settings as a model of session monitoring during toileting.

Acknowledgments

We would like to acknowledge all of the staff and participants at the locations evaluated in this investigation. We are grateful to them for their participation and helpfulness through this process.

Conflicts of Interest

At the time of the publication, VK, PS, THL, and GK were employees at Toi Labs, receiving salary, stock options, and benefits. PC worked in a consultancy role and received stock options. JG has no conflicts of interest to report at the time of this publication.

Multimedia Appendix 1
Exit survey on user acceptance and satisfaction.
[DOCX File , 15 KB - aging_v7i1e50856_app1.docx ]

Multimedia Appendix 2
Example chart of a facility’s toileting logging timing throughout a single day.
[DOCX File , 72 KB - aging_v7i1e50856_app2.docx ]

References


Abbreviations

- AUC: area under the curve
- HIPAA: Health Insurance Portability and Accountability Act
- ROC: receiver operating characteristic
- RWE: real-world evidence
- TL: TrueLoo

©Jordan Glenn, Parmoon Sarmadi, Paul Cristman, Gabrielle Kim, Ting-Hsuan Lin, Vikram Kashyap. Originally published in JMIR Aging (https://aging.jmir.org), 27.05.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Efficacy of COMPAs, an App Designed to Support Communication Between Persons Living With Dementia in Long-Term Care Settings and Their Caregivers: Mixed Methods Implementation Study

Ana Inés Ansaldo, PhD; Michèle Masson-Trottier, MPO, PhD; Barbara Delacourt, MSc, CCO; Jade Dubuc, MSc; Catherine Dubé, MSc

Laboratoire de Plasticité cérébrale, Communication et Vieillissement, Centre de recherche de l'Institut Universitaire de gériatrie de Montréal, Université de Montréal, Montréal, QC, Canada

John Hopkins Hospital, John Hopkins University, Baltimore, MD, United States

Corresponding Author:
Ana Inés Analdo, PhD
Laboratoire de Plasticité cérébrale, Communication et Vieillissement
Centre de recherche de l'Institut Universitaire de gériatrie de Montréal
Université de Montréal
4565 Queen Mary Road
Montréal, QC, H3W1W5
Canada
Phone: 1 5143403540 ext 3933
Email: ana.ines.ansaldo@umontreal.ca

Abstract

Background: Persons living with dementia experience autonomy loss and require caregiver support on a daily basis. Dementia involves a gradual decline in communication skills, leading to fewer interactions and isolation for both people living with dementia and their caregivers, negatively impacting the quality of life for both members of the dyad. The resulting stress and burden on caregivers make them particularly susceptible to burnout.

Objective: This study aims to examine the efficacy of Communication Proches Aidants (COMPAs), an app designed following the principles of person-centered and emotional communication, which is intended to improve well-being in persons living with dementia and caregivers and reduce caregiver burden.

Methods: In this implementation study, volunteer caregivers in 2 long-term care facilities (n=17) were trained in using COMPAs and strategies to improve communication with persons living with dementia. Qualitative and quantitative analyses, semistructured interviews, and questionnaires were completed before and after 8 weeks of intervention with COMPAs.

Results: Semistructured interviews revealed that all caregivers perceived a positive impact following COMPAs interventions, namely, improved quality of communication and quality of life among persons living with dementia and caregivers. Improved quality of life was also supported by a statistically significant reduction in the General Health Questionnaire-12 scores (caregivers who improved: 9/17, 53%; z=2.537; P=0.01). COMPAs interventions were also associated with a statistically significant increased feeling of personal accomplishment (caregivers improved: 11/17, 65%; t15=2.430; P=.03; d=0.61 [medium effect size]).

Conclusions: COMPAs intervention improved well-being in persons living with dementia and their caregivers by developing person-centered communication within the dyad, increasing empathy, and reducing burden in caregivers although most caregivers were unfamiliar with technology. The results hold promise for COMPAs interventions in long-term care settings. Larger group-controlled studies with different populations, in different contexts, and at different stages of dementia will provide a clearer picture of the benefits of COMPAs interventions.

(JMIR Aging 2024;7:e47565) doi:10.2196/47565

KEYWORDS

dementia; communication; caregivers; technology; burden; mixed methods design; quality of life; mobile phone; tablet
**Introduction**

**Background**

Dementia is a progressive or chronic syndrome, affecting memory, reasoning, orientation, understanding, calculation, learning ability, language, and judgment; it represents a greater impairment of cognitive function than might be expected, while being the main cause of disability and dependence among older people [1]. Dementia is the consequence of diverse diseases, the most common being Alzheimer disease [1].

According to the World Health Organization, 55 million individuals worldwide live with dementia, making it one of the main causes of disability and social deprivation among older adults [2]. These figures are set to rise over the next few years, with an estimated 78 million people expected to be living with dementia by 2030. The report states that support for the care of people living with dementia and help for caregivers need to be stepped up urgently. As dementia progresses, it has an impact on the autonomy of the person affected. At the advanced stage, persons living with dementia can no longer live at home because they are no longer able to carry out everyday tasks (eg, dressing, eating, and washing), and this is often difficult to manage for those around them.

**Person-Centered Communication and Dementia**

Dementia is characterized by a progressive deterioration of cognitive functions and particularly affects language [3]. As the disease progresses, persons living with dementia will experience increasing communication deficits that impact all aspects of life. Frequent communication breakdowns lead to feelings of frustration that can trigger reactive behaviors known as behavioral and psychological symptoms of dementia (BPSD) [4]. Communication breakdowns complexify care, increase caregiver burden, and decrease the quality of life (QoL) for persons living with dementia and caregivers [5,6]. In the later stages of dementia, sustaining a simple communicative exchange (eg, greetings or short conversation between the caregiver and the person living with dementia) is practically impossible. Because they are unable to express their needs or understand others, persons living with dementia often express frustration [7] and generally require more attention than other older adults, which also contributes to increasing the burden on caregivers and decreasing QoL [8,9].

The literature shows that communication is a key component of quality care [10-13] and a core component of person-centered care, promoting positive social interactions around topics of the life story of persons living with dementia [5]. Person-centered care, considered state of the art in dementia care [14], recognizes the individual as a person and aims to respond to the individual’s feelings, preferences, and needs [15,16]. Furthermore, person-centered care precludes perceiving the person living with dementia exclusively as a person with illness; and such perception contributes to cognitive decline, adds to communication difficulties, and contributes to depersonalization [13]. A person-centered nondirective approach considers the person’s lifestyle, culture, and history, including their likes and dislikes, preferences, and interests, while always considering the person’s point of view [5].

In sum, communication breakdowns contribute to depersonalization and weakening of person-centered care [17], requiring continuous adaptation from the caregiver. In contrast, person-centered care develops leadership in caregivers, prompting management changes toward a more personalized philosophy of care.

**Impact of Communication on Care**

Communication deficits in the context of dementia have negative impacts on several aspects of the caregiver and resident relation within a long-term care (LTC) setting. More specifically, the progressive nature of the illness entails frequent communication breakdowns, which generate frustration in both the caregiver and the person living with dementia (referred to as the dyad) [5,6]. Indeed, most persons living with dementia show signs of frustration when they cannot understand a conversation or make themselves understood [7,18]. Frustration increases emotional tension, which in turn contributes to the caregiver’s burden [9]. Poor-quality interactions also increase the risk of agitation and apathy in persons living with dementia [8]. While the quality of communication within the dyad is known to modulate caregiver burden [9], it also affects person-centered care, both of which are essential to QoL [19,20]. Greater burden and a higher prevalence of anxiety disorders are observed in caregivers working with persons living with dementia [21,22]. Burden is described along 2 dimensions: objective burden, which refers to the degree of dependence of the person living with dementia and the presence of BPSD, and subjective burden, which is associated with the physical, social, and emotional dimensions of caring, as well as the resources available to the caregiver [23]. A systematic review by Queluz et al [24] grouped professional caregiver needs into 3 main themes: emotional health, formal or informal help received from third parties, and need for information about dementia and associated care.

Several variables related to the physical and social environment in which communication occurs can create living conditions that promote or hamper QoL in caregivers and persons living with dementia [20]. These include the quality of caregivers’ engagement in care (ie, a positive attitude), enjoyable communication as reflected by personalized exchanges [25], and social activities [26]. In addition, the progression in communication deficits often results in avoidance of communication within the dyad, a factor that contributes to accelerating cognitive decline and triggering BPSD, which are particularly disruptive in LTC settings [27]; it also impacts caregivers’ QoL [28]. A review by Scott-Cawiezell et al [29] on 995 staff members has shown that improvements are required to achieve open, accurate, and timely communication in nursing homes. More specifically, according to McCormack et al [30], while some staff members know of residents’ preferences, this information is not routinely communicated to all staff members in a facility. By sharing information about communication topics and strategies facilitating person-centered communication with each resident, caregivers may become more efficient in providing care and less exposed to communication breakdowns and the resulting increase in their burden. Moreover, according to Kolanowski et al [31], available communication systems do not consider the time and resource constraints of nursing homes.

https://aging.jmir.org/2024/1/e47565

JMIR Aging 2024 | vol. 7 | e47565 | p.98

(page number not for citation purposes)
There is a strong consensus on the need to empower caregivers and give them strategies so that they can optimize communication with persons living with dementia [28]. Particularly in the context of LTC residences, the evidence shows that adopting a person-centered care approach significantly influences quality of care and the QoL of both caregivers and persons living with dementia, by improving care compliance and reducing caregivers’ burden [26,28]. Moreover, the literature shows that personalized, emotionally relevant contents facilitate person-centered care, preserve personhood, and prevent dehumanization and isolation of persons living with dementia [5]. Sharing complete information with the persons living with dementia and their families, ensuring participation of persons living with dementia and their families in decision-making, and securing the collaboration of persons living with dementia in policy and program development are among the key elements of person-centered care [32]. Thus, while communicating with persons living with dementia in the provision of care, it is important to consider the unique life history, feelings, cultural background, values, and preferences of each person. This consideration is a challenge for a caregiver in an LTC setting who meets a person living with dementia in the advanced stages when they are unable to tell their life story. Communication tools that consider this important aspect of person-centered care while fitting into the reality of LTC settings (including tight schedules, resources, and constraints, together with administrative investment in nursing leadership) are therefore required to achieve beneficial changes.

Sustaining person-centered care and communication between persons living with dementia and their caregivers, especially in LTC settings, requires adapted communication tools. It was with this purpose that our team designed Communication Proches Aidants (COMPAs; it also refers to the compass, the instrument that orients sailors in troubled seas).

COMPAs was designed to support person-centered communication between persons with severe communication impairments and their caregivers, professional or informal. It is based on the concept of person-centered communication, as reflected by its personalized audiovisual content in line with the life trajectory, preferences, interests, and culture of the persons living with dementia. Through coviewing, it engages the person living with dementia and the caregiver in a form of dialogue beyond words as they share verbal and nonverbal expressions of joy and well-being. Unlike the purely transactional communication that is characteristic of basic care [17], COMPAs puts the person living with dementia at the center of communication, thereby providing a unique framework for person-centered care in the context of dementia.

In sum, the person-centered care approach acknowledges the person living with dementia as their own person. Communication is an essential tool in the implementation of this approach and the maintenance of personhood [33]. Persons living with dementia have trouble expressing their needs, which often leads to the perception that they have no awareness, and their interactions lack mutuality [33]. However, persons living with dementia need meaningful interactions [34], and caregivers need adequate communication tools to meet the care and social needs of persons living with dementia. Unfortunately, caregivers lack sufficient communication training and adapted tools to support social communication in order to overcome the communication deficits of persons living with dementia. Furthermore, LTC settings are environments where there are multiple, often changing, caregivers who revolve around the persons living with dementia. Technology could play a crucial role in care for persons living with dementia. According to Koo and Vizer [35], technology facilitates daily activities, maintains social interactions, supports autobiographical memory, and promotes leisure activities, all while allowing storing and monitoring the clinical status of individuals. Overall, technology could improve the QoL of persons living with dementia, reducing the BPSD and burden on caregivers [35].

Technology and Communication

Overview

Although the evidence shows that technology is relevant for promoting social interactions [35-37], there is a need for rigorous studies on the use of digital tablets in the context of persons living with dementia and their caregivers. Specifically, a scoping review on this topic points to the benefits of technology for intergenerational communication [38], in particular, by means of suggested conversation topics related to the life trajectory of the person living with dementia, which generate positive emotions; however, the review acknowledged that little attention is paid to higher-level needs, such as self-esteem and preserving personhood [35]. Regarding the use of tablets, evidence shows that persons living with dementia may enjoy using tablets at all stages of dementia [38,39]. Hung et al [40] has pointed out the utility of tablets in the context of one-on-one, small-group, and large-group activities, thereby facilitating relationship building and resident engagement and helping caregivers gain better knowledge of the interests and abilities of the persons living with dementia. Furthermore, a recent review [37] has identified a series of apps that could potentially prevent and overcome communication barriers. However, none of these apps were designed to promote person-centered communication or were tested in an LTC setting. Finally, a recent scoping review [35] on the use of technology in LTC has shown some positive impacts on behavior engagement and mood in LTC residents. Importantly, the authors highlight that this finding is not specific to persons living with dementia, who are generally excluded from such studies, while pointing out the importance of examining the impact of technology use specifically in this population [35].

In sum, technology offers some promise for supporting communication in persons with dementia. However, there is limited empirical research on the use of technology to support communication between caregivers and residents in LTC residences [35,41], and only 1 study examined the caregivers’ perspectives on the ability of mobile apps to support caregiver-resident communication [36]. The consensus in the literature is that more research on the use of communication apps involving persons living with dementia in LTC settings is imperative [35]. Moreover, there is a need for evidence-based apps specifically designed to promote person-centered communication in cases of advanced dementia when the possibilities of verbal communication are very limited.
**COMPAs App**

COMPAs is an evidence-based app available on digital tablets to facilitate portability and participation of persons living with dementia. It is designed to support person-centered care between persons living with dementia and their caregivers. This app integrates knowledge from proven therapies [6,15,42-46] and has the advantage of combining these concepts in a single medium [47]. COMPAs is a secure platform that collects photos, music, and videos that have marked the life of the persons living with dementia. It was designed by speech language pathologists (SLPs) and media experts and offers an intuitive environment to promote positive communicative moments; its content is fully customizable to reflect the relevant life events of each person living with dementia and is constantly adaptable as the person’s dementia evolves.

With COMPAs, caregivers gain secure access to personalized libraries of audiovisual materials that are selected according to residents’ personal preferences, cultural background, and life history. Specifically, photographs, music, and videos can be uploaded to a secure space through the tablet or the COMPAs website, allowing families to remotely add content to enrich their loved one’s COMPAs space. Apart from the content provided by the family, the COMPAs artificial intelligence module provides access to personalized internet content related to the person’s life history (eg, places where the person lived, favorite sports, animals, and hobbies). Caregivers can also add music pieces (from music libraries available on the tablet or through a Spotify account) as well as excerpts from movies or shows available on YouTube, while family caregivers can safely add personal videos. A “like” allows users to highlight particularly meaningful content. Before closing COMPAs each time, the caregiver is given a short multiple-choice questionnaire, which gathers a more personal perspective on the COMPAs session, while providing a means to follow up its effects and share relevant information on session outcomes with the members of the team. Coviewing sessions with COMPAs have been shown to facilitate person-centered care, the gold standard in dementia care. Studies conducted during the COVID-19 pandemic showed that COMPAs sessions triggered emotional communication, characterized by shared verbal and nonverbal exchanges related to positive emotions, while increasing social engagement between persons living with dementia and caregivers [47]. Finally, by adding likes to specific music, videos, or photographs in the residents’ personalized space and writing comments, caregivers can share information with the team about the best communication strategies and topics to sustain person-centered communication.

The application was initially developed in French; however, because it uses little written language and is very intuitive, neither language impairments nor language barriers prevent its use by speakers of other languages (see Figure 1 showing the app’s interface).

The rationale behind COMPAs lies in the person-centered care approach and emotional communication. Its personalized, meaningful content reduces the impact of communication deficits by encouraging nonverbal and emotional communication. Regulation of positive emotions has been shown to optimize care in the context of dementia, while favoring trust and promoting well-being for the person living with dementia [48] and the caregiver [49].

Pilot work by our team has shown that COMPAs allows caregivers and LTC residents to enjoy moments of person-centered interaction, breaking the vicious circle of noncommunication, thereby reducing both residents’ isolation and caregivers’ burden. COMPAs’s theoretical background and our pilot findings offer some promise regarding the app’s potential to support person-centered communication between persons living in LTC residences and their caregivers.

**Purpose of the Study**

In light of these pilot findings and considering the need for evidence-based technology tools to support person-centered communication between persons living with dementia and their caregivers in LTC settings [35,41], the purpose of this study was to test the use of COMPAs in LTC settings. Specifically, we implemented COMPAs in the context of LTC daily routines and measured its effects on the communication between

![Figure 1. Communication Proches Aidants (COMPAs) interface.](https://aging.jmir.org/2024/1/e47565/figure1.png)
residents and caregivers, caregivers’ burden, and the QoL of persons living with dementia and caregivers.

In line with the literature, and considering the rationale underlying COMPAs, it was expected that interventions with COMPAs would achieve the following:

1. Improve the quality of communication between the person living with dementia and the caregiver, as measured through improvements in questionnaire scores and semistructured interviews.
2. Enhance the QoL of the persons living with dementia, as measured through improvements in questionnaire scores and semistructured interviews.
3. Reduce the burden on caregivers, as measured through improvements in questionnaire scores and semistructured interviews.

In addition, we anticipated that COMPAs would be adapted to the LTC environment. This had been assessed through participant adherence and satisfaction with the use of COMPAs in this study.

**Methods**

**Study Design**

This study used a pretest-posttest experimental design, with a COMPAs intervention administered over 8 weeks. The total duration of the study, including recruitment and assessments, was 14 weeks. The timeline depicted in Figure 2 shows the various stages of the COMPAs study.

**Participant Selection Process, Inclusion, and Exclusion Criteria**

**Caregivers**

In total, 17 caregivers were recruited. The inclusion criteria were being a caregiver at the Paul-Bruchési LTC Center in Montreal, Québec, Canada, or Saint-Victor LTC center in Amiens, France, and caring for a person living with dementia who presented communication impairments as described in their chart or perceived by the caregiver. There were no exclusion criteria for caregivers. However, 1 participant from the Saint-Victor center was subsequently excluded from the study analyses, as only quantitative data were available for this participant; consequently, 16 caregivers completed the study. Most of the caregivers included in this study were women (14/17, 82%). Caregivers were between the ages of 24 and 57 years and had between 1 and 29 years of work experience in the health sector. Before this study, none of the caregivers had used a digital tablet, and only some used a smartphone (4/17, 22%) at work, although they were all familiar with these technologies since more than half of them used a digital tablet (10/17, 59%) or a smartphone (15/17, 88%) at home. Of the 22 caregivers who attended the information session and were not included in the study, 3 (14%) were not able to participate because they were transferred to another workplace, 1 (4%) refused to participate for personal reasons, and 1 (4%) left on parental leave.

**Residents**

The inclusion criteria were having an assigned caregiver enrolled in the project and experiencing communication difficulties, as identified by a caregiver, in the context of a diagnosis of major neurocognitive impairment, whether isolated or in combination with hearing and visual loss, or other conditions that can challenge communication, including a linguistic barrier. To make the samples as representative of the LTC population as possible, there were no exclusion criteria for persons living with dementia. A total of 17 residents participated in the study; they were aged between 61 and 96 years. Most of the residents (11/17, 65%) had a diagnosis of dementia (mixed dementia: n=6, 35%; Alzheimer disease: n=3, 18%; vascular dementia: n=1, 6%; severe dementia: n=1, 6%; and Lewy body disease: n=1, 6%). Other diagnoses were hippocampal atrophy (n=1, 6%), generalized anxiety (n=1, 6%), and cancer (n=1, 6%). Most of the residents (n=12, 71%) included in this study were women and still had some ability to express themselves verbally in isolated words or short utterances, with fluctuating comprehension of short sentences (n=15, 88%), including due to hearing limitations (n=6, 36%).

**Recruitment Process**

The recruitment process for caregivers was on a voluntary basis. Specifically, the project was presented by the research assistant (RA) and the laboratory director (Ana Inés Ansaldo) during a staff meeting. The purpose of the presentation was to stimulate interest in the study and to introduce COMPAs to the staff members. Staff members were asked to contact the head of the LTC unit to express their interest in participating in the study. Each caregiver identified a resident with whom they wanted to improve communication. The research team then asked the residents if they were interested in participating.

**Ethical Considerations**

Ethics approval was obtained from the Centre de recherche de l’Institut universitaire de gériatrie de Montréal (CRIUGM) Ethics
Committee (approval number CER-18-19-14), and informed written consent was obtained from the residents or their representatives in cases of incapacity. Caregivers were invited to sign consent forms with the RA after they had expressed interest to participate in the study.

Pre-Experimental Phase

Information Session

The first week of the study was dedicated to describing the project’s purpose, the procedures, the measurement tools, and COMPAs’s characteristics and use.

Training Sessions on COMPAs Use

Caregivers received two 30-minute training sessions on COMPAs during the daily planned team meetings. Facilitated by an SLP or a trained RA, the training session focused on COMPAs’s rationale and principles and a demonstration of its use by the trainer. There was a hands-on practice session at the end of the training.

Pre-Experimental Measures

The measures administered to caregivers included questionnaires investigating QoL (General Health Questionnaire-12 [GHQ-12] items) [50] and the burden at work (Maslach Burnout Inventory [MBI]) [51]. The RA also administered measures related to residents, including an overview of residents’ communication deficits (Grille d’évaluation des difficultés de communication dans la démence [GCOM]) [52] and a QoL questionnaire (Qualité de vie dans la démence de type Alzheimer [QDV-DTA]) [53]. Information on the residents’ and caregivers’ age, sex, and other sociodemographic data was also collected, as was information on residents’ neurocognitive disorders from their medical charts.

The GHQ-12 was used to acquire data on the caregivers’ general QoL. It includes 12 questions, scored on a 4-point Likert scale where 1=not at all, 2=not more than usual, 3=a little less than usual, and 4=a lot more than usual. In the standard scoring system, scores of 1 or 2 are given 0 point, and scores of 3 or 4 are given 1 point. The overall score is the sum of 0 and 1 points. If the sum is higher than 2, it is considered problematic.

The MBI measures caregiver burden at work. It comprises (1) an emotional exhaustion score (9 questions; scores of <17, 18-29, and >30 denote low, moderate, and high emotional exhaustion levels, respectively); (2) a depersonalization score (5 questions; scores of <5, 6-11, and >12 denote low, moderate, and high empathy loss, respectively); and (3) a personal achievement score (8 questions; scores >40, 34-39, and <33 denote low, moderate, and high achievement levels, respectively). Caregivers were asked to rate their own scores on this test.

The GCOM measures the severity of communication difficulties. The caregiver is asked to score communication behaviors for each resident (eg, “The person has word-finding difficulties”) on a Likert scale ranging from 1 (always) to 4 (never); it is also possible to select “does not speak enough for me to judge.” All scores are added up to give a final communication score. The QDV-DTA measures residents’ QoL. It consists of 13 questions, with higher scores indicating better QoL.

Setting Up Personalized COMPAs Libraries

Before the intervention, the RA completed a personal history with the residents’ representatives. The caregivers, along with an SLP or a trained RA, created a personalized communication space for each resident by adding significant personal content such as personal photos, images, videos, and music selected based on their life history questionnaire and the information provided by the residents’ representatives.

Experimental Phase

Caregivers used COMPAs for 8 weeks in the context of their daily LTC routine (eg, family visits; recreation time; birthday celebrations; and situations triggering reactive behaviors, such as personal care or specific interventions). COMPAs sessions could be very short (2-5 minutes) or longer (up to 20 minutes). They consisted of coviewing sessions (ie, resident and caregiver) during which personalized material in the resident’s library was presented by the caregiver, with the purpose of eliciting positive emotions and triggering verbal and nonverbal exchanges within the dyad, according to the principles learned during the training sessions.

Caregivers were instructed to ensure that residents always used their hearing and visual aids during COMPAs sessions. Furthermore, following caregivers’ comments, we made adaptations such as using personal hearing amplifiers or Bluetooth speakers to improve listening to the music. In addition, since COMPAs is a person-centered approach, we encouraged caregivers to focus on the person’s strengths and therefore used more photos and videos with people with hearing loss and more music and audio clips with those with visual loss.

Caregivers could modify the number and duration of COMPAs sessions according to what was possible for them (eg, workload and residents’ status), as long as they respected a minimum of 15 minutes a day, at a time they considered appropriate. During the first 2 weeks of the intervention, with the purpose of facilitating COMPAs use, the SLP and the trained RA provided direct support to the caregivers on the telephone, by email, or during visits to the LTC facility. All caregiver shifts were covered so that everyone had the chance to ask questions. Indirect support (by telephone and email) was available throughout the 8-week duration of the intervention.

Postexperimental Phase and Measures

Following the 8 weeks of COMPAs intervention, the same measures administered before the experiment were administered by the RA to the caregivers and residents. Semistructured interviews with the participating caregivers were also completed: individual, semistructured 20- to 40-minute interviews were conducted with the caregivers in person, and on the Zoom (Zoom Video Communications Inc) platform. Only the interviewer and the caregiver were present at the meeting. The questions came from an interview guide developed by the last author (CD), based on interview guides from previous studies [54,55]. The interview consisted of open-ended questions and began with a very general question (“Could you please talk..."
about your experience with COMPAs in the last few weeks?”). Uncertainties arising from participants’ answers were elucidated with follow-up questions. At the end, the interviewer also asked the caregiver if they wished to provide any additional information. Overall, 17 interviews were conducted with the 17 participants.

### Data Analysis Plan

#### Quantitative Analyses

Data analyses include quantitative and qualitative approaches. Primary outcome measures are scores on the GHQ-12 [50] and the BMI [51] with caregivers and scores on the GCOM [52] and QDV-DTA [53] with persons living with dementia. The 4 outcome measures were used as dependent variables, and the 2 measurement points before and after the COMPAs intervention were considered as independent variables. The results were entered in paired sample 2-tailed $t$ tests with an $\alpha$ of .05 to define significance. The paired-sample $t$ test allows to control for individual variables that potentially affect outcome measures. Furthermore, paired-sample $t$ tests are suitable to analyze data from small samples, as in this study. When the assumptions of the paired-sample $t$ test were not fulfilled (ie, difference scores were not normally distributed and within-participant variability was not consistent), a nonparametric alternative, the Wilcoxon signed-rank test, was used. For all measures, outliers were detected by inspection; criteria to exclude data from analyses were set at values >1.5 box lengths from the edge of the box plot.

#### Qualitative Analyses

The qualitative measures involved a qualitative content analysis approach [56], as described by Intissar and Rabeb [57] and Vallée et al [58]. To analyze data from semistructured interviews with caregivers, MAXQDA software (VERBI GmbH), a qualitative analysis program for discourse content analyses was used [59]. The interviews were audio recorded and manually transcribed verbatim by BD and JD, and the interjudge validity was assessed, followed by thematic coding. The corpus of the interviews was read multiple times separately by 2 authors to achieve a full understanding of the data. Transcripts were coded individually by 2 authors (CD and JD), and multiple meetings took place to reach intercoder reliability. Each author separately generated codes, following which they met to discuss them and reached consensus on the coding tree, which was further discussed with the last author (CD) of this manuscript. Quotations presented in this paper were translated from French into English. Caregivers were assigned numbers from P1 to P17.

### Results

#### Quantitative Findings

#### Caregivers

Outlier values were detected in the difference scores for the GHQ-12, the MBI-depersonalization score and the MBI-personal achievement score. All data points >1.5 box lengths, each associated with a different participant, were removed from this specific analysis. The differences between the MBI-emotional exhaustion score, MBI-depersonalization score, and MBI-personal achievement score scores were normally distributed ($P=.39$, $P=.29$, and $P=.74$, respectively). Thus, paired-sample $t$ tests were used to analyze these differences. The results are presented in Table 1. In contrast, the GHQ-12 scores were not normally distributed ($P=.02$), and therefore, a Wilcoxon signed-rank test was used to analyze these differences. The results are presented in Table 2. Data are presented as mean (SD), unless otherwise stated.
Table 1. Individual scores and mean (SD) values on subsections of the Maslach Burnout Inventory.

<table>
<thead>
<tr>
<th>ID (scores)</th>
<th>Emotional exhaustion</th>
<th>Depersonalization</th>
<th>Personal achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Postest</td>
<td>Difference</td>
</tr>
<tr>
<td>CA³</td>
<td>43</td>
<td>33</td>
<td>−10</td>
</tr>
<tr>
<td>CA4</td>
<td>18</td>
<td>21</td>
<td>3</td>
</tr>
<tr>
<td>CA6</td>
<td>16</td>
<td>19</td>
<td>3</td>
</tr>
<tr>
<td>CA7</td>
<td>20</td>
<td>4</td>
<td>−16</td>
</tr>
<tr>
<td>CA10</td>
<td>9</td>
<td>28</td>
<td>19</td>
</tr>
<tr>
<td>CA11</td>
<td>17</td>
<td>15</td>
<td>−2</td>
</tr>
<tr>
<td>CA15</td>
<td>10</td>
<td>7</td>
<td>−3</td>
</tr>
<tr>
<td>CA16</td>
<td>14</td>
<td>6</td>
<td>−8</td>
</tr>
<tr>
<td>FR¹</td>
<td>14</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>FR2</td>
<td>11</td>
<td>1</td>
<td>−10</td>
</tr>
<tr>
<td>FR4</td>
<td>20</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>FR5</td>
<td>11</td>
<td>12</td>
<td>1</td>
</tr>
<tr>
<td>FR6</td>
<td>20</td>
<td>9</td>
<td>−11</td>
</tr>
<tr>
<td>FR7</td>
<td>54</td>
<td>51</td>
<td>−3</td>
</tr>
<tr>
<td>FR8</td>
<td>13</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>FR9</td>
<td>36</td>
<td>33</td>
<td>−3</td>
</tr>
<tr>
<td>FR10</td>
<td>2</td>
<td>15</td>
<td>13</td>
</tr>
</tbody>
</table>

Values

<table>
<thead>
<tr>
<th>Group, mean (SD)</th>
<th>Emotional exhaustion</th>
<th>Depersonalization</th>
<th>Personal achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Postest</td>
<td>Difference</td>
</tr>
<tr>
<td>Group, mean (SD)</td>
<td>19.29</td>
<td>18.00</td>
<td>−1.29 (8.68)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Outlier, mean (SD)</th>
<th>Emotional exhaustion</th>
<th>Depersonalization</th>
<th>Personal achievement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pretest</td>
<td>Postest</td>
<td>Difference</td>
</tr>
<tr>
<td>Outlier, mean (SD)</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

ᵃCA: caregivers from the Canadian site.
ᵇFR: caregivers from the French site.
ᶜNot applicable.
Table 2. Individual scores for caregivers on the General Health Questionnaire-12 (GHQ-12) and pre- and postintervention mean scores with all participants.

<table>
<thead>
<tr>
<th>ID</th>
<th>Pretest</th>
<th>Posttest</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA³₅</td>
<td>4</td>
<td>3</td>
<td>–1</td>
</tr>
<tr>
<td>CA4</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CA6</td>
<td>3</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>CA7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CA10</td>
<td>3</td>
<td>0</td>
<td>–3</td>
</tr>
<tr>
<td>CA11</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CA15</td>
<td>1</td>
<td>0</td>
<td>–1</td>
</tr>
<tr>
<td>CA16</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FR¹</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FR2</td>
<td>2</td>
<td>0</td>
<td>–2</td>
</tr>
<tr>
<td>FR4</td>
<td>1</td>
<td>0</td>
<td>–1</td>
</tr>
<tr>
<td>FR5</td>
<td>5</td>
<td>2</td>
<td>–3</td>
</tr>
<tr>
<td>FR7</td>
<td>8</td>
<td>3</td>
<td>–5</td>
</tr>
<tr>
<td>FR8</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>FR9</td>
<td>2</td>
<td>1</td>
<td>–1</td>
</tr>
<tr>
<td>FR10</td>
<td>3</td>
<td>0</td>
<td>–3</td>
</tr>
</tbody>
</table>

Values

<table>
<thead>
<tr>
<th>Group, mean (SD)</th>
<th>Mean without outliers (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.88 (2.26)</td>
<td>1.5 (1.67)</td>
</tr>
<tr>
<td>0.76 (1.20)</td>
<td>0.62 (1.08)</td>
</tr>
<tr>
<td>–1.11 (1.57)</td>
<td>–0.87 (1.25)</td>
</tr>
</tbody>
</table>

³CA: caregivers from the Canadian site.
¹FR: caregivers from the French site.

To determine whether the use of COMPAs by caregivers in LTC settings influenced their scores on the GHQ-12 questionnaire, the Wilcoxon signed-rank test was used to measure QoL scores. The difference scores were approximately symmetrically distributed, as assessed by a histogram with a superimposed normal curve. Of the 17 caregivers who participated in this study, 9 (53%) showed a decrease in score difference, 7 (41%) showed a tied score, and 1 (6%) showed an increase in score difference. There was a statistically significant change (mean –0.88, SD 1.25; median –0.5, IQR 2.50) between GHQ-12 scores at T3 (mean 1.50, SD 1.67; median 1, IQR 3.0) and at T14 (mean 0.63, SD 1.09; median 0, IQR 1.5; z=–2.54; P=.01). At T14, 1 (6%) caregiver’s score had worsened, 9 (53%) had improved, and 7 (41%) remained the same.

A 2-tailed paired-sample t test was used to determine whether COMPAs influenced caregivers’ scores on the MBI-emotional exhaustion scale, MBI-depersonalization scale, and MBI-personal achievement scale. No significant change from T3 (mean 19.29, SD 13.20) to T14 (mean 18.00, SD 12.67; t₁₅=–0.62; P=.55) was found on the MBI-emotional exhaustion scale. Specifically, at T3, 10 (59%) caregivers scored a low level of emotional exhaustion, 4 (24%) scored a moderate level, and 3 (18%) scored an elevated level. At T14, 10 (59%) caregivers were at a low level, 4 (24%) at a moderate level, and 3 (18%) at an elevated level; scores had worsened for 7 (41%) caregivers, improved for 9 (53%), and remained the same for 1 (6%).

There was also no significant change on the MBI-depersonalization scale (T3: mean 7.75, SD 4.84; T14: mean 6.88, SD 5.06; t₁₅=–1.01; P=.33). At T3, 5 (29%) caregivers had a low level of empathy loss, 9 (53%) had a moderate level, and 3 (18%) had a high level. At T14, 8 (47%) caregivers scored a low level of empathy loss, 5 (29%) scored a moderate level, and 4 (24%) scored a high level; 9 (53%) caregivers had improved, 5 (29%) had worsened, and 3 (18%) remained at the same level.

Regarding personal achievement, according to the MBI-personal achievement scale, there was a significant change following COMPAs use (T3: mean 37.24, SD 10.21; T14: mean 38.71, SD 9.2; t₁₅=2.43; P=.03; d=0.61 for a medium effect size). Specifically, at T3, a total of 9 (53%) caregivers scored a high level of personal achievement, 4 (24%) scored a moderate level, and 4 (24%) scored a low level. At T14, 9 (53%) caregivers...
scored a high level of personal achievement, 4 (24%) scored a moderate level, and 4 (24%) scored a low level; 5 (29%) caregivers had worsened, 11 (65%) had improved, and 1 (6%) remained the same.

Residents
A paired-sample t test was used to determine whether there was a statistically significant mean difference between residents’ overall scores on the QDV-DTA and the GCOM before and after the use of COMPAs by caregivers in the LTC setting. There were no outliers in the data or in the QDV-DTA differences or in the GCOM differences, as assessed by the inspection of a box plot for values >1.5 box lengths from the edge of the box.

The assumption of normality was not violated, as assessed by the Shapiro-Wilk test for the QDV-DTA or the GCOM (P=.78 and P=.78, respectively). The results did not reveal any significant change in the overall score on the QDV-DTA (QDV-DTA: preintervention mean 33.42, SD 3.92; postintervention mean 34.58, SD 5.42; t₁₁=0.84; P=.42) or the GCOM (GCOM: preintervention mean 16.42, SD 10.08; postintervention mean 20.67, SD 12.92; t₁₁=1.89; P=.09). However, 2-tailed t tests revealed a significant worsening for the following questions in the GCOM: “They tend to repeat something that someone just said” (P=.03); “They use filler words (‘thing,’ ‘whatchamacallit’) instead of precise words (‘pencil,’ ‘balloon’)” (P=.04).

Although the QoL questionnaire and the GCOM did not show a significant change, the semistructured interviews with all caregivers (n=17) revealed that they felt COMPAs had a positive impact on the lives of persons living with dementia (please refer to the Qualitative Findings section). Some caregivers described these positive impacts as positive emotions during nonverbal communication revealed by a positive facial expression in the person living with dementia or the simple fact that the person living with dementia started dancing while using COMPAs.

Qualitative Findings
The qualitative analyses of semistructured interviews with the caregivers regarding the effects of COMPAs use with persons living with dementia in the LTC setting highlighted 3 main themes: capacity of COMPAs to elicit positive emotions, decrease in caregiver burden, and versatility of COMPAs.

Eliciting Positive Emotions
The interviews highlighted COMPAs’s capacity to elicit positive emotions in persons living with dementia and caregivers. This increase in positive emotions led to an increase in feelings of joy, pleasure, and happiness.

Persons With Dementia
Caregivers (16/16, 100%) reported on COMPAs’s capacity to elicit positive emotions in persons living with dementia, who had expressed well-being and the pleasure they got from their COMPAs session both verbally and nonverbally via changes in their facial expression (e.g., smiles and eye contact):

This nonverbal person's face would light up. [P6]

Almost all caregivers discussed the positive effects of personalized content on persons living with dementia: how pictures or music from their past can evoke positive memories (15/17, 88%) and positive emotions (16/17, 94%). Persons living with dementia enjoyed reminiscing and sharing former moments of their life with their caregivers:

COMPAs calms them, and does them good too, because it reminds them of memories, good memories, there are pictures of their kids, their pet, their house. [P14]

I would say cheerfulness. They were happy to see the pictures, listen to their favorite music. [P17]

Caregivers
Caregivers (14/17, 82%) reported that using COMPAs gave them joy, pleasure, and overall positive effects on their daily lives. They said that they looked forward to using COMPAs with their patients and appreciated these moments in their week:

To me, it is my moment of pleasure, when I am working. [P5]

It was good for them but for me too because I felt their well-being. [P17]

Some of the caregivers (14/17, 82%) said that simply seeing the enjoyment of the persons living with dementia gave them pleasure too:

Yes, I would see that it brought them joy, so it brought me pleasure too. [P14]

Caregivers also reported on how they appreciated the effects of the app on the persons living with dementia. They valued COMPAs’s capacity to contribute to the well-being of persons living with dementia:

It’s great, because we give the resident the chance not to feel lost, and without this tool, you can’t really do it. [P10]

Decrease in Caregiver Burden
The analysis highlighted how COMPAs gave caregivers a solution to deal with their everyday struggles, helped them feel empowered, and resulted in better bonds with the residents.

COMPAs was used as a tool and, in some cases, as an excellent way out of difficult situations involving persons with dementia, such as opposition, disorientation, and apathy. Caregivers (12/17, 71%) saw COMPAs as a solution that worked with residents who had required more attention. They were grateful to have an effective solution in these types of cases:

Interviewer: Did you have the impression of having a solution?

P13: Yes, exactly, now I have a solution.
P2: But, with some residents, the device also helps us to do the tasks. There are some cases that are more difficult, but with the device, it improves our interaction with certain residents a bit.

The implementation of COMPAs also contributed to caregivers’ feelings of personal achievement in the workplace. Caregivers (15/17, 88%) felt more useful and believed that they made a
real difference in the lives of the persons living with dementia, as they could go beyond providing primary care. Caregivers enjoyed learning about the residents, their lives, and their personal tastes, and they felt empowered by having an additional clinical role:

It lifts you up, in your work...you are not only there to help them with comfort care or to feed them. [P10]
I can say it adds something good to the atmosphere, it adds more...How can I say this...I could say particularly with PWDs, it makes us want...with people who communicate less. It is like us; it makes us want to reach out to those people. [P6]

Caregivers (11/17, 65%) discussed how COMPAs enabled the development of more personalized relationships with persons living with dementia. They felt closer to the residents and more interested in them. Spending more time with persons living with dementia helped caregivers to create a bond and spend quality time with them despite the communication difficulties:

P9: ...just spending fifteen minutes with them, it was a moment of joy and relaxation. Because I was also learning plenty of things.
Interviewer: Do you mean that it provided you with joy to spending time with them?
P9: Yes, exactly. We never take enough time; we do not take the time to talk to them.

Caregivers (11/17, 65%) also said that COMPAs allowed them to have better interactions with the residents. They talked about better-quality exchanges, and generally enhanced communication, including communication by the persons living with dementia themselves:

We can communicate with the device. [P14]
Even their speech is more fluent. [P10]

Versatility of COMPAs
Caregivers highlighted the flexibility of COMPAs, as it could be used for different reasons, in different settings, and for different durations.

COMPAs was implemented in different ways by different caregivers. They could adapt it to their working conditions, and over time, they incorporated the tool into their daily routine. Some caregivers used it while providing grooming care:

And there are moments, like grooming care, that are a bit stressful, I would put his music on, and we sang, we danced in front of the mirror, and we giggled. [P13]

Some used it in a group setting, while others used it individually. Some caregivers had a fixed time in their day dedicated to COMPAs, while others used it at different times depending on the situation:

A big asset of COMPAs, is that in fact, we can use it at any time of day. [P11]

The duration of a COMPAs session also varied between caregivers, ranging from 5 minutes to around 20 minutes:

When I have five minutes, ten minutes, I would take the iPad, go to the room, and we listened together. Sometimes when I have more time, I stay longer. [P4]

Challenges for Caregivers
Caregivers reported some issues during the implementation of COMPAs. Lack of time, technological issues, and the responsibility for or availability of the device were mentioned as challenges.

Lack of Time
Close to half of the caregivers (8/17, 47%) stated that they lacked the time to use COMPAs. They commented that it was not always easy to take the time to conduct a COMPAs session because of workload or when time permitted, the resident might be unavailable:

The evening shift, it is hard to find the time to enjoy it. [P3]
It was hard in the mornings with grooming care: there is too much work to do it properly. However, we would do, I would do one in the morning from time to time. [P15]

Technological Issues
The participants encountered some technological issues during the implementation. Caregivers (3/17, 18%) discussed how the bugs could disrupt their sessions, making the residents lose interest in the content presented:

The videos did not work. I would have liked to do it with Mrs. B, watch videos, but it was not working. [P17]

Responsibility for and Availability of the Device
Caregivers (6/17, 35%) raised the issue of being responsible for an iPad. Being responsible for a valuable object was a concern for them. In other cases, the iPad was locked, and a nurse had to make it available to the caregiver (5/17, 29%). Their busy schedules made it difficult for them to access the device when they needed it.

Challenges for Residents
A few residents (6/17, 35%) faced some problems while using COMPAs. Confusion, negative emotions, and disinterest were mentioned as challenges for residents.

Caregivers reported that some of their residents considered the app to be an intrusion; they did not understand how their personal information came to be in the attendant’s hand:

P10: People like Mrs. G, this dementia, well, for her, it’s not positive, because it’s difficult, she starts questioning. She wonders what is going on.
Interviewer: You think Mrs. G., it makes her wary?
P10: Mrs. G., she did it once and it was very hard to do it again, because she takes it as an intrusion.

Others were troubled by not being able to recognize the pictures shown to them. Negative emotions could also be elicited by the content:

https://aging.jmir.org/2024/1/e47565 JMIR Aging 2024 | vol. 7 | e47565 (page number not for citation purposes)
A little bit of melancholy at times. [P12]

Caregivers (4/17, 24%) reported that some residents were disengaged from COMPAs. This disinterest was related to the device, the redundancy of the content, or the resident’s attitude:

Well, there were some residents who weren’t even slightly interested. [P13]

Some residents weren’t interested in watching the screen or were troubled by the screen. [P13]

Discussion

Principal Findings

The purpose of this study was to implement and validate the effect of COMPAs, an app designed to elicit positive emotions triggering communication between persons living with dementia and their caregivers in an LTC setting.

Using a combination of quantitative and qualitative methods, the main results of the study validate our hypotheses. Specifically, the qualitative results from the semistructured interviews show that COMPAs improved person-centered communication between caregivers and persons living with dementia: its use resulted in more verbal and nonverbal exchanges in different contexts (eg, personal care and dedicated time). In particular, caregivers reported an improvement in the quality of exchanges and a more personal care relationship. Moreover, the use of COMPAs was associated with an improvement in QoL for both persons living with dementia and caregivers. The caregivers reported that COMPAs elicited positive emotions in persons living with dementia, contributing to emotional communication and helping the caregivers see the person living with dementia beyond the illness. In so doing, COMPAs supported person-centered care and communication between persons living with dementia and their caregivers. In addition, statistically significant results were observed in the form of increased caregiver empowerment, as reflected by the accomplishment score in the MBI. Caregivers also described COMPAs as a solution that helped them create opportunities to develop meaningful bonds with persons living with dementia, easing the caregiver communication burden. Finally, COMPAs was deemed well suited to the LTC context, particularly due to its versatility. These results were observed although 40% (7/17) of the caregivers were not accustomed to using an iPad, which provides evidence for the versatility of COMPAs in empowering even caregivers with limited technological literacy.

Quality of Communication Between Persons Living With Dementia and Caregivers

The GCOM showed some deterioration in specific oral expression components for persons living with dementia, which is expected in the context of progressive conditions. Interestingly, the COMPAs intervention was associated with stable general communication skills in residents. This may be a result of the GCOM’s poor sensitivity to the communication patterns characterizing advanced neurocognitive disease, or it might reflect the benefits of daily stimulation with COMPAs in reducing morbidity, despite the progressive nature of neurocognitive disease [7]. The results on the GCOM also highlighted the positive nonverbal communication markers of well-being and positive emotions that COMPAs induced in residents, including smiling; raising eyebrows; touching the caregiver’s hand while coviewing; smiling, dancing, or singing to personalized music; and laughing with the caregiver. These findings highlight the app’s efficacy in promoting person-centered communication between LTC residents with dementia and their caregivers. They also attest to the benefits of COMPAs training, which increases caregivers’ awareness of nonverbal and emotional aspects of communication. These findings are in line with those of previous work showing that integrating communication strategies into care and using elements of a patient’s life story in informal discussions enhance meaningful communication between caregivers and LTC residents [6]. Moreover, the results of the semistructured interviews showed that following the COMPAs trial, caregivers focused more on nonverbal and emotional person-centered communication and less on verbal and transactional communication. These results reveal the importance of combining a good tool with suitable training in order to promote awareness of all dimensions of communication and the potential facilitators and barriers [31].

Studies conducted during the COVID-19 pandemic also found that COMPAs had positive effects on communication between caregivers and persons living with dementia in an LTC setting, even during periods of extreme isolation [47]. Specifically, caregivers reported that residents showed increased engagement, as opposed to apathy, together with verbal and nonverbal expressions of joy, well-being, and calmness while using COMPAs, despite major public health restrictions and the use of personal protective equipment [47].

Residents’ and Caregivers’ QoL

Caregivers’ QoL

The results on the GHQ-12 showed that caregivers’ QoL increased following COMPAs use, and so did their sense of personal accomplishment (measured by the MBI-personal achievement scale). More specifically, an improvement in the feeling of personal accomplishment was observed in the level of energy that the caregiver felt when working closely with the resident. This is in line with the findings of previous studies showing that significant communication between persons living with dementia and caregivers is associated with a better QoL [60,61].

The analyses of semistructured interviews show that caregivers described COMPAs-supported interactions with residents as pleasant times, moments of relaxation, and even as therapeutic for them. Thus, caregivers saw COMPAs as a solution to their struggles with persons living with dementia: a quick and effective way out of the challenges they faced in managing difficult behaviors (eg, apathy and agitation) and engage in more natural communication. It was probably this factor that led to the association between reduced caregiver burden and COMPAs use. Similar findings were reported in previous work showing the relationship between caregiver burden and the quality of communication [9,62]. Furthermore, there was a decrease in the score of the MBI-depersonalization scale for over half of the participating caregivers (9/17, 53%). This may have been related to the personalized COMPAs content, which helps...
caregivers appreciate the person beyond the disease and become aware of the individual traits of the persons living with dementia including their history, culture, tastes, and preferences, that is, the opposite of depersonalization.

The caregivers expressed their satisfaction with knowing more about the person they were caring for, spending more time with them, getting to know their life story better, and seeing the residents happier in this context. Indeed, personalized content is shown to be relevant in facilitating communication among persons living with dementia [63]. Thus, improving the quality of communication had positive effects on the dyad and helped to establish an empathic relationship. As a result, caregivers felt valued and satisfied with their work:

> It lifts you up in your job: you’re not there just to help them with their comfort care or feed them.

Using COMPAs empowered the caregivers, and this is probably a key reason for reduced caregivers’ burden.

**Residents’ QoL**

The QDV-DTA scores did not show any significant changes in the residents’ QoL. However, caregivers reported that COMPAs triggered positive expressions in the residents, demonstrated by their nonverbal communication. They viewed COMPAs as a tool helping residents change their routine, remember positive times, break out of their isolation, and feel well. They considered COMPAs to be a valuable tool to support meaningful communication, thus supporting social engagement in persons with dementia. During the semistructured interviews, caregivers mentioned COMPAs’s ability to support a meaningful activity, which meant persons living with dementia were involved in stimulating activities. In light of the literature, residents’ QoL is promoted by social contacts, a good relationship with the caregiver, and the caregiver’s involvement in providing care [26,28]. These factors are also associated with better self-esteem in persons living with dementia, an essential component of their well-being [5] and dignity [17]. Furthermore, it is essential to address the socialization needs of persons living with dementia and provide person-centered care [64]. In line with this literature and considering the verbal and nonverbal manifestations of well-being in residents documented by caregivers in the semistructured interviews, the results of this study prove the relevance of COMPAs for communication and QoL. They illustrate the value of orienting communication around emotional content linked to the residents’ own life trajectory, which improves interactions in the dyad, and in turn promotes positive relationships [16,33].

**Burden on Caregivers**

QoL at work refers to various factors such as satisfaction, mental health, and stress level [65]; these 3 factors contribute to caregiver burden. More specifically, stress arises when caregivers fear not having the necessary resources to face the physical and psychological challenges that they may encounter in geriatric care. The reduction in caregiver burden and increased personal achievement found in this study may be related to several factors. One is the fact that caregivers received a training session on communication strategies and barriers in LTC. Thus, the literature shows that training on strategies for communicating with persons living with dementia improves caregivers’ communication skills and is associated with a decrease in their burden [61].

Another factor that may have contributed to reducing caregiver burden is reflected by the results on the MBI and by the thematic analyses concerning the stress and frustration associated with communication barriers [9], all of which were reduced in this project. Specifically, the quantitative results on the MBI following 8 weeks of COMPAs use showed a significant reduction in the burden score, concurrently with a significant improvement in the personal achievement score and a reduction in feelings of tiredness. Hence, COMPAs was a resource for caregivers, allowing them to feel more accomplished and less exhausted at work. These effects are also illustrated by the semistructured interviews; caregivers reported that adding COMPAs to their daily routine did not result in work overload.

In fact, caregivers noticed the positive effect of COMPAs on the residents and wanted to continue using the app, although they occasionally mentioned not having time to use COMPAs to their satisfaction or to add material to the libraries. The key will be to find more time in everyday life situations for caregivers to use COMPAs and to simplify the addition of personalized material. We are currently working on these 2 elements to meet these needs.

In sum, the results on the MBI and semistructured interviews show that COMPAs reduced caregiver burden, a factor that is associated with more relaxed care, which in turn further reduces their burden [66,67]. Furthermore, the emotional component of COMPAs interventions contributed to the expression of empathy, which is also known to reduce caregivers’ burden.

**Strengths and Limitations of the Study**

To our knowledge, this is the first study to explore the implementation of a communication-based app in the context of LTC settings. The qualitative results of this study demonstrate COMPAs’s effectiveness in positively influencing the lives of persons living with dementia and caregivers in an efficient and timely manner. It should be noted that COMPAs’s positive effects on caregivers might perhaps be influenced by a selection bias; they volunteered for the study, and therefore, they might have a positive bias toward the method and would not necessarily be representative of the target population. However, according to the LTC administrators, the caregivers’ demographic profile in the sample was representative of the vast majority of caregivers in Québec and Canada, mostly immigrant women aged between 20 and 60 years. Following the launch of the project, more caregivers saw the benefits and told the research team that they wish they had enrolled. Future studies could explore COMPAs’s effects and adherence in a larger community of caregivers as a function of cultural background, age, gender, and technological literacy, among other things.

Finally, we acknowledge that the statistical significance of the quantitative results is limited. This may be a consequence of the small sample size and the diversity of residents’ clinical profiles. Hence, although we find medium effect sizes, these results should be considered as only a tendency. Future studies...
with larger samples of participants are required to confirm these results and test the generalizability of these findings in broader populations, including family caregivers of persons living with dementia in LTC residences or other LTC populations without dementia but with severe communication impairments following stroke, or severe sensory impairments in the context of behavioral disturbances, or major psychiatric disorders.

COMPAs interventions proved to be suitable for implementation in LTC residences, as caregivers could choose the duration, time, and modality of their sessions with persons living with dementia. Caregivers could adapt the app to their working conditions, making COMPAs a versatile tool that can be modulated to the users’ needs.

The literature shows that technology is underused to support the communication between caregivers and persons living with dementia in LTC settings [37], and no available evidence was found to support the use of apps for such communication. This study provides evidence that COMPAs is suitable to support person-centered care in the caregiver-LTC resident dyad. To our knowledge, this is the first app that supports communication through a person-centered care approach and that is fully customizable to the person living with dementia.

Although the information gathered from caregivers in the questionnaires and the semistructured interviews is an effective way to understand COMPAs’s effects on persons living with dementia, we acknowledge that the perspectives of persons living with dementia were not included. Including them in the research team entails ethical and logistical challenges [68,69]. To be in line with the person-centered care approach, future studies should include their point of view [32]. The research team created the personalized libraries, which saved time for the caregivers, who were overloaded with everyday work, but persons with dementia could be more involved in choosing the material. This will be done in future studies.

In a context where human resources and time are limited, it will be important to develop efficient ways to create user-friendly personalized libraries. Future work will focus on improving the COMPAs interface by adding artificial intelligence modules that will assist in more efficiently creating sophisticated personalized libraries.

Conclusions
COMPAs proved to be effective in improving communication between caregivers and residents while reducing the burden on caregivers and improving both groups’ QoL.

The evidence shows that COMPAs facilitates person-centered communication. Positive emotions generated in residents resonate in caregivers, stimulating empathy and well-being in the dyad. This state of shared well-being promotes social engagement, defocuses attention from impairments and disabilities, and fosters exchanges between the dyad and what they share. The impressive gains in relevant outcome measures obtained with residents and caregivers underscore the relevance of COMPAs in LTC settings. Large-scale studies are necessary to validate the observed tendencies and optimize COMPAs’s potential benefits in persons with living dementia in LTC settings and their caregivers, while examining its use with other susceptible populations presenting communication deficits. Studies could also explore the barriers to technology use in caregivers and persons living with dementia and ways to overcome them; they must also consider the ethical issues related to technology use with susceptible populations, including privacy and security, to identify best practices for safe implementation of technology in dementia care.

Acknowledgments
The authors are grateful to all the residents and caregivers who participated in the study and to the administrators and members of the research team whose cooperation made this study possible.

This work was supported by the Spark Program of the Center for Aging Brain Health Innovation, powered by Baycrest.

Conflicts of Interest
None declared.

References


8. Bourgeois MS. “Where is my wife and when am i going home?” The challenge of communicating with persons with dementia. Alzheimers Care Today 2002;3(2):132-144 [FREE full text]


52. Murphy J, Gray CM, Cox S. Communication and dementia: how talking mats can help people with dementia to express themselves. University of Glasgow. 2007. URL: http://eprints.gla.ac.uk/61498/ [accessed 2024-04-29]


59. MAXQDA. VERBI Software. URL: https://www.maxqda.com/about [accessed 2024-04-29]


63. Hird N, Osaki T, Ghosh S, Palaniappan SK, Maeda K. Empowering personalization for digital cognitive stimulation to support communication with people with dementia: pilot intervention study as a prelude to AI development. JMIR Form Res 2024 Jan 16;8:e51732 [FREE Full text] [doi: 10.2196/51732] [Medline: 38227357]


Abbreviations

- **BPSD**: behavioral and psychological symptoms of dementia
- **COMPAs**: Communication Proches Aidants
- **GCOM**: grille d’évaluation des difficultés de communication dans la démence
- **GHQ-12**: General Health Questionnaire-12
- **LTC**: long-term care
- **MBI**: Maslach Burnout Inventory
- **QDV-DTA**: qualité de vie dans la démence de type Alzheimer
- **QoL**: quality of life
- **RA**: research assistant
- **SLP**: speech language pathologist

https://aging.jmir.org/2024/1/e47565
Efficacy of COMPAs, an App Designed to Support Communication Between Persons Living With Dementia in Long-Term Care Settings and Their Caregivers: Mixed Methods Implementation Study

JMIR Aging 2024;7:e47565
URL: https://aging.jmir.org/2024/1/e47565
doi:10.2196/47565
PMID:38963691

©Ana Inés Ansaldo, Michèle Masson-Trottier, Barbara Delacourt, Jade Dubuc, Catherine Dubé. Originally published in JMIR Aging (https://aging.jmir.org), 04.07.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
The Frailty Trajectory’s Additional Edge Over the Frailty Index: Retrospective Cohort Study of Veterans With Heart Failure

Javad Razjouyan1,2,3, PhD; Ariela R Orkaby4,5, MD; Molly J Horstman1,2, MD; Parag Goyal6, MD; Orna Intrator7,8, MD; Aanand D Naik2,9,10, MD

Corresponding Author:
Javad Razjouyan, PhD

JMIR Aging 2024;7:e56345 doi:10.2196/56345

KEYWORDS
gerontology; geriatric; geriatrics; older adult; older adults; elder; elderly; older person; older people; ageing; aging; frailty; frailty index; frailty trajectory; frail; weak; weakness; heart failure; HF; cardiovascular disease; CVD; congestive heart failure; CHF; myocardial infarction; MI; unstable angina; angina; cardiac arrest; atherosclerosis; cardiology; cardiac; cardiologist; cardiologists

Introduction
Individuals with heart failure (HF) have a high burden of health care utilization, costs, and morbidity in the year following hospitalization for an acute HF exacerbation. Frailty, which has been described as increased vulnerability to adverse events, is common among those with HF and increases with age [1]. Health systems worldwide are integrating automated tools within electronic health records to measure frailty. However, using longitudinal data to measure frailty and better predict outcomes among those with HF has rarely been considered [2-5]. We sought to evaluate the predictive value of adding longitudinal data to a standard frailty index (FI) and evaluate predictions of 1-year outcomes in patients with HF.

Methods
Study Design
This was a retrospective cohort study that used national Veterans Health Administration (VA) data. Veterans aged ≥50 years with an index hospital admission for HF from 2016 to 2019 were included. We excluded veterans with <2 primary care visits in the 3 years before their date of admission to indicate regular use of VA care. We included those with documentation of ejection fraction. We used the validated VA FI, which captures 31 deficits in health based on International Classification of Diseases, Tenth Revision, and Current Procedural Terminology codes [6]. We estimated the FI for each preceding year, without overlap. We fit a linear line to 3 calculated FIs for each year prior to the index date of admission and reported the slope and intercept individually. This method provided a 3-year longitudinal estimate of frailty at admission. We used 1-year all-cause mortality following the index date of admission as the primary outcome. We reported the area under the curve (AUC) for predicting outcomes, using logistic regression. We estimated two AUCs: (1) FI at the time of admission (AUC_FI) and (2) FI at time of admission plus slope and intercept (AUC_FI+trajectory). Changes in the AUCs were reported as the percentage of improvement (ΔAUC = 100% × [AUC_FI+trajectory – AUC_FI]/AUC_FI). We recursively calculated the AUCs and ΔAUC by including patients whose FIs at admission were <0.1 and, at each step, increased the FI level by 0.01 to 0.4.

Ethical Considerations
The study protocol was approved by the Research & Development Committee of the Michael E. DeBakey VA Medical Center and Baylor College of Medicine Institutional Review Board (institutional review board number: H-464220).

Results
In total, 54,774 veterans were included (age: mean 73.3, SD 10.1 y; BMI: mean 30.1, SD 7.5 kg/m²; male: n=53,899, 98.4%; White: n=30,406, 55.5%; Table 1). Figure 1 shows the AUC_FI and AUC_FI+trajectory across the distribution of frailty ranges, from prefrail (FI: 0.1-0.2) to frail; an FI of 0.2 is equivalent to an
accumulation of 7 deficits among 31 variables, and the $\Delta AUC$ is also displayed. For all veterans across all FI thresholds, the AUC improved by at least 4.1% when adding the FT to the FI. The highest $\Delta AUC$ (24%) was observed for FIs of 0.13 to 0.16, and it decreased to $\leq$10% for FIs of $\geq$0.2.

Table. Characteristics of patients (N=54,774) with an index admission to the Veterans Health Administration for heart failure from January 1, 2016, to January 1, 2020.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admit year 2016, n (%)</td>
<td>12,875 (23.5)</td>
</tr>
<tr>
<td>Admit year 2017, n (%)</td>
<td>13,585 (24.8)</td>
</tr>
<tr>
<td>Admit year 2018, n (%)</td>
<td>14,082 (25.7)</td>
</tr>
<tr>
<td>Admit year 2019, n (%)</td>
<td>14,232 (26)</td>
</tr>
<tr>
<td>Age (y), mean (SD)</td>
<td>73.3 (10.1)</td>
</tr>
<tr>
<td>&lt;65, n (%)</td>
<td>9776 (17.8)</td>
</tr>
<tr>
<td>65 - 75, n (%)</td>
<td>22,772 (41.6)</td>
</tr>
<tr>
<td>$\geq$85, n (%)</td>
<td>22,226 (40.6)</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td>53,899 (98.4)</td>
</tr>
<tr>
<td>Male</td>
<td>53,899 (98.4)</td>
</tr>
<tr>
<td>Female</td>
<td>875 (1.6)</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td>30,406 (55.5)</td>
</tr>
<tr>
<td>White</td>
<td>30,406 (55.5)</td>
</tr>
<tr>
<td>Black</td>
<td>9340 (17.1)</td>
</tr>
<tr>
<td>Other$^a$</td>
<td>15,028 (27.4)</td>
</tr>
<tr>
<td>Hispanic ethnicity, n (%)</td>
<td>2093 (3.8)</td>
</tr>
<tr>
<td>BMI (kg/m$^2$), mean (SD)</td>
<td>30.1 (7.5)</td>
</tr>
<tr>
<td>$\geq$30, n (%)</td>
<td>24,352 (44.5)</td>
</tr>
<tr>
<td>Frailty status (frailty index), mean (SD)</td>
<td>0.35 (0.11)</td>
</tr>
<tr>
<td>Robust ($&lt;0.1$), n (%)$^b$</td>
<td>297 (0.5)</td>
</tr>
<tr>
<td>Prefrail (0.1 - 0.2), n (%)$^b$</td>
<td>5715 (10.5)</td>
</tr>
<tr>
<td>Frail ($&gt;0.2$), n (%)$^b$</td>
<td>48,762 (89)</td>
</tr>
<tr>
<td>All-cause mortality, n (%)</td>
<td>2848 (5.2)</td>
</tr>
<tr>
<td>30-day mortality</td>
<td>2848 (5.2)</td>
</tr>
<tr>
<td>1-year mortality</td>
<td>14,460 (26.4)</td>
</tr>
<tr>
<td>All-time mortality</td>
<td>37,027 (67.6)</td>
</tr>
<tr>
<td>Time to death (mo), median (IQR)</td>
<td>18.2 (5.6-36.4)</td>
</tr>
<tr>
<td>HFrEF$^c$, n (%)</td>
<td>27,223 (49.7)</td>
</tr>
<tr>
<td>HFmEF$^d$, n (%)</td>
<td>4546 (8.3)</td>
</tr>
<tr>
<td>HFpEF$^e$, n (%)</td>
<td>23,005 (42.0)</td>
</tr>
<tr>
<td>Living in a CLC$^f$, n (%)</td>
<td>1808 (3.3)</td>
</tr>
</tbody>
</table>

$^a$“Other” includes Asian, American Indian or Alaska Native, Native Hawaiian or other Pacific Islander, and unknown.

$^b$Standardized frailty status cut points drawn from validated studies [6].

$^c$HFrEF: heart failure with reduced ejection fraction of <$40%$.

$^d$HFmEF: heart failure with modified reduced ejection fraction of 40%-50%.

$^e$HFpEF: heart failure with preserved ejection fraction of $>50%$.

$^f$CLC: community living center.
Figure 1. AUCs for patients who were admitted, for the first time, to the Veterans Health Administration for heart failure from January 1, 2016, to January 1, 2020, and had an FI of 0.1-0.4 (as shown on the x-axis in increments of 0.01). We compared the AUCs of FIs (in blue; AUC_{FI}) versus the AUCs of FIs and FTs combined (in orange; AUC_{FI+FT}). The percentage of improvement in AUCs resulting from the addition of the FT to the FI was reported in black (\Delta AUC) and calculated by using the following formula: \Delta AUC = \frac{(AUC_{FI+FT} - AUC_{FI})}{AUC_{FI}} \times 100. AUC: area under the curve; FI: frailty index; FT: frailty trajectory.

Discussion

In a national cohort of veterans who were admitted to the VA for HF, the addition of longitudinal FT data resulted in a clinically significant (up to 24%) improvement in 1-year mortality prediction when compared to a standard FI alone among patients in the prefrail range. In contrast, we observed a modest (at least 4.1%) improvement in 1-year mortality prediction in the overall population. Enhancing AUC prediction for patients in the prefrail range is clinically important, as interventions that mitigate frailty may be most impactful in this population [7]. Patients with prefrailty may benefit from interventions (e.g., cardiac rehabilitation) that improve frailty status and cardiovascular outcomes [1]. These findings enrich our understanding of the importance of FT in patients at lower FI levels, and a previous study compared the importance of FIs to that of FTs alone [5]. These results may not generalize to nonveteran populations. The sample was predominately male but did include a diverse population in terms of race, ethnicity, and geographic distribution. In summary, methods for
calculating frailty provide useful predictions of adverse outcomes among adults with HF. The addition of longitudinal frailty data improves predictions for patients with HF and prefrailty. These findings aid clinician and health system decision-making, as this population benefits most from interventions that slow or prevent frailty progression, and suggest that longitudinal data for modeling FT provide additional evidence for tailoring interventions to patients with HF who may benefit most from tailored interventions.

Acknowledgments
JR is supported by seed funding from Baylor College of Medicine, Houston, TX, United States; the Center for Innovations in Quality, Effectiveness and Safety (CIN 13–413), Michael E. DeBakey VA Medical Center, Houston, TX, United States; and National Institutes of Health (NIH), National Heart, Lung, and Blood Institute (NHLBI) K25 funding (1K25HL152006-01). ARO is supported by VA Clinical Science Research & Development Career Development Award Level 2 (award IK2-CX001800). MJH is supported by VA HSR CDA-2 award 1K2HX003163-01A2.

Conflicts of Interest
None declared.

References

Abbreviations
AUC: area under the curve
FI: frailty index
FT: frailty trajectory
HF: heart failure
VA: Veterans Health Administration

© Javad Razjouyan, Ariela Orkaby, Molly Horstman, Parag Goyal, Orna Intrator, Aanand D Naik. Originally published in JMIR Aging (https://aging.jmir.org), 27.6.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction
in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Hospital Staff Perspectives on the Drivers and Challenges in Implementing a Virtual Rehabilitation Ward: Qualitative Study

Leanne Greene¹, PhD; Miia Rahja¹, PhD; Kate Laver¹, PhD; Vun Vun Wong², MBBS; Chris Leung², BPhty; Maria Crotty², PhD

Corresponding Author: Leanne Greene, PhD

Abstract

Background: Over the past decade, the adoption of virtual wards has surged. Virtual wards aim to prevent unnecessary hospital admissions, expedite home discharge, and enhance patient satisfaction, which are particularly beneficial for the older adult population who faces risks associated with hospitalization. Consequently, substantial investments are being made in virtual rehabilitation wards (VRWs), despite evidence of varying levels of success in their implementation. However, the facilitators and barriers experienced by virtual ward staff for the rapid implementation of these innovative care models remain poorly understood.

Objective: This paper presents insights from hospital staff working on an Australian VRW in response to the growing demand for programs aimed at preventing hospital admissions. We explore staff’s perspectives on the facilitators and barriers of the VRW, shedding light on service setup and delivery.

Methods: Qualitative interviews were conducted with 21 VRW staff using the Nonadoption, Abandonment, Scale-up, Spread, and Sustainability (NASSS) framework. The analysis of data was performed using framework analysis and the 7 domains of the NASSS framework.

Results: The results were mapped onto the 7 domains of the NASSS framework. (1) Condition: Managing certain conditions, especially those involving comorbidities and sociocultural factors, can be challenging. (2) Technology: The VRW demonstrated suitability for technologically engaged patients without cognitive impairment, offering advantages in clinical decision-making through remote monitoring and video calls. However, interoperability issues and equipment malfunctions caused staff frustration, highlighting the importance of promptly addressing technical challenges. (3) Value proposition: The VRW empowered patients to choose their care location, extending access to care for rural communities and enabling home-based treatment for older adults. (4) Adopters and (5) organizations: Despite these benefits, the cultural shift from in-person to remote treatment introduced uncertainties in workflows, professional responsibilities, resource allocation, and intake processes. (6) Wider system and (7) embedding: As the service continues to develop to address gaps in hospital capacity, it is imperative to prioritize ongoing adaptation. This includes refining the process of smoothly transferring patients back to the hospital, addressing technical aspects, ensuring seamless continuity of care, and thoughtfully considering how the burden of care may shift to patients and their families.

Conclusions: In this qualitative study exploring health care staff’s experience of an innovative VRW, we identified several drivers and challenges to implementation and acceptability. The findings have implications for future services considering implementing VRWs for older adults in terms of service setup and delivery. Future work will focus on assessing patient and carer experiences of the VRW.

(JMIR Aging 2024;7:e54774) doi:10.2196/54774

KEYWORDS

gerontology; geriatric; geriatrics; older adult; older adults; elder; elderly; older person; older people; ageing; aging; aged; telerehabilitation; rehabilitation; rehab; workflow; hospitalization; health services accessibility; accessibility; clinical decision-making; equipment failure; telemedicine; telehealth; tele-medicine; tele-health; virtual care; virtual health; virtual medicine; remote consultation; telephone consultation; video consultation; remote consultations; telephone consultations; video consultations; personnel; hospital

https://aging.jmir.org/2024/1/e54774

JMIR Aging 2024 | vol. 7 | e54774 | p.120
Implementation rates of virtual wards have increased over the last decade, mainly driven by technological advancements and the COVID-19 pandemic [1-5]. The UK National Health Service (NHS) report that virtual wards, including hospital-at-home services, are a safe and efficient substitute for inpatient care that is facilitated by technology [2,6,7]. Virtual wards offer at-home acute care in the form of monitoring and treatment to individuals who would otherwise require a hospital bed, using a flexible combination of remote and in-person services [2,6].

As virtual models of care are relatively novel, there can be ambiguity around terminology [2,8]. In the NHS model, virtual wards for older individuals function akin to hospital-at-home services, primarily delivering care through face-to-face interactions [8].

Virtual wards aim to prevent avoidable hospital admissions, facilitate early discharge home, and increase patient satisfaction [3,6,9]. For the older adult population, hospital admissions carry potential risks, including deconditioning, delirium, and hospital-acquired infections [10], so opting for home-based treatments may be beneficial [11]. Consequently, significant investments are being directed toward the expansion of virtual care models in health care systems, such as the UK NHS, encompassing patients with frailty [12,13]. However, economic assessments of virtual models often fail to meet quality criteria, leading to varying estimated cost savings [2].

The implementation of virtual care models has been hindered by issues such as nonadoption, abandonment, and difficulties with scaling up, particularly if the model requires significant changes to the broader care system [14-16]. There is a paucity of research investigating the sustainability of virtual models [15], particularly virtual wards [2]. Despite substantial policy-level discussions and modest proof-of-concept studies, virtual health care models are seldom mainstreamed [17-19]. The success or failure of implementing innovative virtual health care models is often attributed to a complex combination of facilitators and barriers, rather than individual factors alone, such as time pressures, infrastructure, unreliable equipment, and staff and service user preferences [20]. Understanding these issues is important as virtual care marks a monumental change in the delivery of health care for older individuals [2].

In practice, virtual wards are often added to existing hospital services as a solution to a bed capacity problem rather than being designed from the ground up as new freestanding services [21-23]. When hospital-based staff are asked to establish a virtual ward and commence providing services through videoconferencing and monitoring, significant shifts in practice are required. The facilitators and barriers experienced by hospital staff for the rapid implementation of these novel care models are not well understood, as evidenced by the paucity of literature in the area. There is also a lack of guidance for the provision of virtual wards, with calls for information on how these new models of care are being implemented [8]. This paper addresses a gap in the existing literature by offering insights into the experiences, perceptions, and attitudes of hospital staff working within a newly established Australian virtual rehabilitation ward (VRW). To our knowledge, no previous studies have delved into this specific research area. The implementation of the VRW was undertaken by the Flinders Medical Centre, part of the Southern Adelaide Local Health Network (SALHN) in South Australia, in response to the increasing demand for programs aimed at preventing hospital admissions. We explored the facilitators and barriers of the VRW from the view of staffs and present reflections for service setup and delivery.

Methods

Design

A multidisciplinary research team (clinicians and academic researchers) conducted this study under a constructivist paradigm [24]. Data were integrated and analyzed using the Nonadoption, Abandonment, Scale-up, Spread, and Sustainability (NASSS) framework [15] to understand staff experiences (Figure 1). Interview questions focused on the 7 domains of the NASSS framework including condition, technology, value proposition, adopters, organizations, wider system, and embedding and adaptation over time. These domains provided an analytical framework for organizing, classifying, and contrasting staff experiences into a rich narrative. The NASSS framework was chosen as it was designed to evaluate technology-supported change projects in health or social care [15] and, therefore, fitted with the aim of our research. Moreover, it has been previously used to evaluate technology-supported health care programs [25-27]. Other frameworks such as the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework [28]; Precede-Proceed model; Dynamic Sustainability framework; or the Practical, Robust Implementation and Sustainability Model (PRISM) [29] were not chosen as they lacked the technology focus of the NASSS framework. We also report our study according to the COREQ (Consolidated Criteria for Reporting Qualitative Research) guidelines [30] to improve the quality and transparency of our work.
Figure 1. The Nonadoption, Abandonment, Scale-up, Spread, and Sustainability (NASSS) framework, first published in 2017, provides a structured approach to examine the factors that impact the adoption, nonadoption, abandonment, spread, scale-up, and sustainability of health care technology (reproduced from Abimbola et al [27], which is published under Creative Commons Attribution 4.0 International License [31]).

Ethical Considerations

Ethics approval was received from the SALHN (Southern Adelaide Clinical Human Research Ethics Committee; 2022/HRE00107). Written informed consent was obtained from all participants, who were informed of their right to withdraw from the study at any time. The data presented in this article have been deidentified. Participants did not receive any compensation.

Setting

South Australia spans a geographical area of 983,482 km$^2$ with a population of 1.8 million [32]. SALHN services a community of approximately 400,000 people in the southern metropolitan area of Adelaide. The population in Southern Adelaide skews toward older age groups compared to other areas of Adelaide and the broader Australian population, with a projected accelerated aging rate [33]. The prevalence of lone-person households, concentrated among older age groups, is rising, and these demographic shifts will likely amplify the demand for health care services [33]. The Flinders Medical Centre is the second-largest tertiary hospital in Adelaide with nearly 600 beds, offering a wide range of medical, surgical, obstetric, and pediatric services.
Virtual Rehabilitation Service

The virtual rehabilitation ward (VRW) provides acute clinical care through rapid assessment and rehabilitation to patients with a range of diagnoses in their own homes. The service provides an alternative to hospital-based rehabilitation and allows patients to be discharged earlier from inpatient wards. A team of multidisciplinary staff (see Table 1), including doctors, nurses, and allied health professionals, work with the individual to achieve rehabilitation goals via a mix of tailored video calls and home visits over a 2-week period with daily clinical reviews. The patient’s medical status is also monitored remotely, for instance, blood pressure, oxygen saturation, temperature, and weight. All equipment is loaned to the patient (eg, iPad with cellular connection via SIM card and monitoring equipment), and training on how to use the equipment is provided by care staff. Home visits can occur as required, and an initial visit is made to set up equipment and provide training.

Table 1. Virtual rehabilitation ward workforce structure.

<table>
<thead>
<tr>
<th>Profession</th>
<th>Full-time equivalent (hours)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical consultant, registrar, and resident medical officer</td>
<td>2.8</td>
</tr>
<tr>
<td>Nursing (clinical and enrolled)</td>
<td>10.5</td>
</tr>
<tr>
<td>Allied health professionals (physiotherapist, occupational therapist, pharmacy, psychology, exercise physiologist, social worker, speech pathologist, dietetics, and allied health assistants)</td>
<td>16</td>
</tr>
</tbody>
</table>

The service is available to individuals over the age of 18 years, but most patients are older adults. Between the commencement of the service in January 2022 and August 2022, a total of 181 (79%) out of 229 patients were aged 65 years or older, with an average age of 73.9 years and a median age of 75.0 years. Local hospital ward staff refer patients to the VRW, and a VRW coordinator assesses the individual’s suitability for the program, usually in person while they are still an inpatient. The VRW accepts a wide range of patients with complex care needs, including individuals recovering from trauma, patients undergoing cancer rehabilitation, or those who require postsurgery care. The service delivers time-limited interventions and monitoring (usually 2 weeks) based on clinical needs. The service runs 7 days a week 24 hours a day, with full staffing between 8 AM and 8 PM and access to an on-call doctor available outside these hours. Patients are provided with a telephone number to contact if their symptoms worsen. The VRW is supported by a contracted external telecommunication provider, which supplements internal SALHN digital health support.

Participant Recruitment

Between July and September 2023, staff previously (ie, rotational junior doctors) or presently employed by the VRW were invited to participate in a semistructured interview exploring the implementation of the VRW. In phenomenological studies, a purposive sampling strategy—in this research, maximum variation sampling—is supported to recruit participants who have experienced the phenomenon under study. We aimed to recruit a diverse range of health care staff (ie, clinical, administration, and information technology) to capture various perspectives on the topic of interest, thereby illuminating diversity and revealing patterns or commonalities in traits across the spectrum [34-38]. Recruitment occurred through the ward managers circulating research information via email and word of mouth. Overall, 21 interviews were conducted, and participant characteristics are presented in Table 2. Data collection was ceased when all the staff who wanted to take part in the study had been interviewed. The number of VRW staff was smaller compared to conventional inpatient rehabilitation wards (Table 1). A total of 9 staff members declined because they were not interested in taking part in research or because they did not have the time to take part in an interview. No participants withdrew from the research after consenting. Each participant was interviewed only once, and no relationship was established prior to study commencement.
Table. Participant demographics (N=21).

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Occupation, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Doctor</td>
<td>6 (29)</td>
</tr>
<tr>
<td>Nurse</td>
<td>3 (14)</td>
</tr>
<tr>
<td>Physiotherapist</td>
<td>3 (14)</td>
</tr>
<tr>
<td>Exercise physiologist</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Occupational therapist</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Information technology officer</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Administrative assistant</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Manager</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Social worker</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Pharmacist</td>
<td>1 (5)</td>
</tr>
<tr>
<td><strong>Age (years), mean (range)</strong></td>
<td>39.1 (25-62)</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>15 (71)</td>
</tr>
<tr>
<td>Male</td>
<td>6 (29)</td>
</tr>
</tbody>
</table>

Qualitative Data Collection and Analysis

We used phenomenology to understand the meaning of the perspectives of staff who worked on the VRW [37]. A researcher independent of the health service conducted the semistructured, audio-recorded interviews (LG). Field notes were made during the interviews to aid reflexivity [39]. A total of 20 interviews were in person and 1 was conducted virtually via Microsoft Teams. To ensure convenience for the staff, in-person interviews were held in a meeting room on the same floor of the hospital as the VRW offices. For the virtual interview, both the researcher and the participant were in their respective homes. Interviews lasted between 45 - 60 minutes and no one else was present besides the participants and researcher. Interview questions and descriptions of the 7 NASSS domains are provided in Multimedia Appendix 1. Questions were tailored for each staff discipline, for example, doctors were asked about remote prescribing while allied health professionals were asked about remote therapy. We did not conduct a pilot test for this piece of research.

Deidentified, audio-recorded interviews were transcribed verbatim by a professional transcription service. One participant requested for their transcript to be returned to them for comment, but no amendments were made. The data were analyzed by 2 coders (LG and MR) using framework analysis [40] to identify the key themes and meanings that emerged from the participants’ descriptions. This method was chosen as it provides a rigorous and transparent approach for researchers to analyze multidisciplinary health research [41]. Moreover, as framework analysis is not aligned with a specific epistemological, philosophical, or theoretical perspective, it adapted well to the use of a preexisting theoretical framework (deductive analysis) while allowing room for revisions with inductive aspects of analysis [40]. In brief, the process of framework analysis involves organizing significant themes and issues into 5 distinct stages: becoming familiar with the data; identifying a thematic framework; indexing; charting; and finally, mapping and interpretation. Multimedia Appendix 2 [15-41] provides a detailed description of the methodology. NVivo 12 (QSR International) was used for coding and indexing the data into the 7 NASSS domains. The information was subsequently condensed and organized into a matrix. Multimedia Appendix 3 provides a description of our coding tree.

Results

Overview

Due to the small size of the VRW staff, where sometimes only 1 member represented each discipline, to ensure anonymity within quoted content, we have grouped them. Physiotherapists, exercise physiologists, occupational therapists, and social workers are grouped under the collective name “Allied Health Professional” and managers and administrative staff are grouped under “Leadership/Admin Team.” Two participants provided feedback on the findings via email and during an informal face-to-face meeting.

In summary, the VRW sought to serve patients with varied health conditions. Challenges arose in managing complex cases such as heart failure and cognitive impairments, impacting staff confidence in virtual care delivery. Despite technological benefits such as remote monitoring, interoperability issues persisted, hindering adoption. The service’s value lay in offering choice and access to care, particularly benefiting rural communities. However, risks included communication challenges with community care teams and less intensive therapy compared to inpatient settings. The transition to virtual care posed workflow and responsibility challenges, highlighting the need for staff training and support. Ambiguity surrounding the service’s identity and referral processes impacted resource allocation and patient expectations. Challenges in patient transfer
and continuity of care were observed, along with resistance to hospital readmission and overcrowded emergency departments. Staff recognized the service’s potential but emphasized the need for specialized planning and ongoing adaptation. Adaptive actions included refining technology and identifying areas for improvement in patient care and service delivery.

**Domain 1: The Condition**

The VRW services patients with several health conditions, including ones that require high-level care (Multimedia Appendix 4). Participants commented on how certain conditions, such as heart failure, fluid retention, and complex wounds, were more challenging to manage using a virtual ward approach in comparison to a traditional inpatient setting.

> We’d have fluid overloaded patients, and you just don’t know how much they’re drinking, or you can’t do the same monitoring as you can in hospital. [Doctor]

Comorbidities (including cognitive impairments, polypharmacy, frailty, disability [eg, limb impairments and tremors], and sensory impairments) were also discussed by all participants as factors impacting how confident the staff felt in engaging patients in the virtual ward service. There was a sense that the service might be better suited for patients with minimal cognitive impairment who are willing to engage with technology.

> In the right population, yes. I think again, if they’re cognitively not good, or they’re really not wanting to engage through technology, then it’s very difficult. But yeah, I think if they’re willing to engage in that, I think it’s no different to being in the room with them. [Allied Health Professional]

The service’s suitability was also influenced by sociocultural factors such as living arrangements. For example, the staff reported relying on carers to assist with virtual sessions or remote monitoring. Therefore, living alone or residing in a care home could potentially pose challenges.

> I think it depends on what support the person has…So if they’ve got someone else there with them who can use the technology that works really well, and there’s definitely no issues there. [Allied Health Professional]

**Domain 2: The Technology**

A lack of interoperability between hardware and software systems was discussed by all 21 participants. There was agreement that the integration between different health care systems, service providers, technology, and security was poor and impeded adoption (Multimedia Appendix 4). This complexity appeared to make it difficult to pinpoint the exact sources of problems when difficulties occurred. Additional challenges came from unforeseen software updates, with the equipment dispersed across patients’ homes (Multimedia Appendix 4).

> It’s like, “Don’t really care whose fault it is, can someone just fix it?” [Allied Health Professional]

The remote monitoring equipment and the capability to make video calls was perceived as advantageous, since it has assisted staff in monitoring patient conditions and facilitating the escalation of care or transferring to another health care service when necessary. This was particularly commented on by doctors, although other staff also made reference to the equipment’s benefits in escalation. Despite staff’s efforts to streamline the process for simplicity, there were reliability and usability concerns, occasionally affecting rapport (Multimedia Appendix 4).

> …it’s not 100% reliable for us, and for us to escalate care of patients based on an unreliable system is difficult to do…We need to have technology that we can rely on, that is safe. [Leadership/Admin Team]

There was, however, recognition that some of these technology failures could stem from a lack of knowledge and education among both staff and patients (Multimedia Appendix 4).

> …maybe we don’t facilitate the education as good as we should. We are out there for what, maybe half an hour. And then, “Here’s your kit, off you go.” [Leadership/Admin Team]

Medtasker, a mobile communication and task management platform, was widely used and commended for its ability to minimize work duplication and reduce the need for excessive phone calls and emails.

> Because the virtual team is sort of here mornings, evenings, there might be people that might have seen the patient a few days and then they’re not the next and so forth. We’re trying to work out the best ways of managing communication. And I think a program like Medtasker helps. [Doctor]

**Domain 3: Value Proposition**

The value of the service was the choice it provided patients on where they received their care (Multimedia Appendix 4). The staff felt that home-based treatment, especially for older individuals, would be a preferred and more comfortable option than staying in hospital. The service was considered empowering and enabled access to care for rural communities who might not be able to access rehabilitation services otherwise (Multimedia Appendix 4).

> I think the gaps that it bridges is amazing and huge, because our country patients, who aren’t able to travel…I see cancer patients who are palliative who really benefit from that ability to be able to connect via video link. [Doctor]

The value of the VRW for patients was also discussed by identifying risks. The opinions about this varied, but in general, there was a sense that the VRW did not pose more risks, just different risks. One concern, particularly with doctors, was that patients often resumed contact with their community care teams (general practitioners [GPs] and community-based medical specialists).

> …managing patients who are still attending their GPs, their specialists in other hospitals in the background, and if you do not have an understanding of what there is happening there and you are involved, I think that is somewhere where some confusion can happen. [Doctor]
Additionally, there were fewer opportunities for physical evaluation, monitoring, and therapy, leading to concerns that clinicians might miss early signs of deterioration. This concern was elaborated upon by 4 of the interviewed doctors, while nurses and allied health professionals also indicated unease about not being able to examine severely ill patients as thoroughly as they typically would on a ward. Allied health professionals discussed patients having less-intensive therapy compared to inpatient wards (Multimedia Appendix 4).

...a lot of clinical signs we aren’t able to pick up by a video link, so we’ll assess them a particular way, but then when the registrar’s gone and done a home visit just because we were worried for whatever reason, we’ve actually found other signs which have triggered escalation of care later, which we weren’t able to identify through video links. [Doctor]

Domain 4: The Adopters

The commencement of the VRW has imposed many changes to staff practices. The interviews delved into the need for a cultural shift and how some participants (and their colleagues) were resistant to adopting the virtual approach. Face-to-face patient interactions were preferred by many, either because they were accustomed to it or because they perceived clinical advantages in such treatments. There was a tendency for conducting initial meetings in person, followed by remote reviews, whenever feasible.

That’s just the way that I’ve nursed for 30 years…I’m not used to trying to do that over a screen. [Nurse]

There appeared to be some uncertainty surrounding workflows and professional responsibilities, leading to concerns that clinical staff might be required to assume technology or administrative duties (Multimedia Appendix 4).

We’re clinical, we’re not IT. [Leadership/Admin Team]

As the service developed, the participants described how they have acquired valuable knowledge about the essential support, training, and resources needed. For instance, the team has recently gained a pharmacist, and this addition has helped alleviate previous prescription medication challenges. Furthermore, due to the unanticipated complex care needs of the referred patients, it has been crucial to have experienced staff to handle the workload effectively. As such, the clinical team has made staffing adjustments to accommodate for the large number of junior staff working within the service (Multimedia Appendix 4). It was highlighted that for a virtual ward service to succeed, the staff should possess strong clinical skills to ensure adaptability and flexibility in care delivery (Multimedia Appendix 4).

...just the more experience and the more training and the more feeding off each other and learning of each other, the more we’re adapting the telehealth. There is a huge education component that’s needed for junior staff and a lot of support for people that aren’t experienced. [Allied Health Professional]

A shift in the burden of care to the patient and their family was also discussed. Staff particularly recognized the essential role that carers play in the patients’ journey during virtual care.

...we heavily use carers and family in this model because not only can they help a lot of the time with setting up the technology and working it out and things like that, they’re quite often present during the reviews and we quite often use them to just be our hands. [Allied Health Professional]

Domain 5: The Organization

There was a sense of ambiguity surrounding the identity of the service both for the staff and patients. The service was set up to provide rehabilitation with a primary focus on functional recovery, but hospital pressures resulted in large numbers of referrals from acute wards, leading to more medically complex referrals than the team had anticipated. This presented a particular challenge for allied health professionals who felt that their roles became almost redundant for certain patients, such as those from oncology wards with newly diagnosed conditions that they were struggling to come to terms with.

We have a lot of palliative care patients coming through lately. And just the questions that they ask, and the difficult conversations, and where to go with rehab. Because it’s not really restorative, it’s more maintenance and optimizing. [Allied Health Professional]

Participants expressed frustration with the intake process, as they believed that referrals should involve patients who are willing and able to use technology as part of their rehabilitation journey.

There needs to be more done in the triage of people’s capacity to use technology…we’ve been getting people who’ve just said to us, “I’m not using the iPad.” [Allied Health Professional]

There was a perception that patients were not accurately informed about the service or that some patients did not fully understand the implications of the term “Virtual Rehabilitation Ward.” Patients often had unrealistic expectations, assuming that rehabilitation at home would be less intensive than in a traditional ward setting (Multimedia Appendix 4).

...they would think that it’s a ward, but then the virtual part of it, sometimes, well some people don’t know what that means. Some people think it’s purely virtual so that there’s no home visiting service. [Leadership/Admin Team]

The importance of resource allocation and the environmental setup was also discussed. Given the rise in the complex patient population requiring more in-person care, vehicle access was often an issue, particularly for nursing and allied health care staff (Multimedia Appendix 4). Open-plan offices and a lack of private spaces for confidential calls were among the challenges discussed (Multimedia Appendix 4). Staff emphasized that the organization and setup of the service were crucial factors in its successful adoption, stressing the need for
a balanced approach that offers a supportive office environment and has enough room and resources to deliver care flexibly.

Because you want to have the team together and interacting like they would on ward in a way because they have that, there’s then that team camaraderie but also that informal sharing of knowledge and teaching each other…but you also need space to be able to operate and talk to patients and hear them. [Doctor]

Domain 6: The Wider Context
Staff felt that the new service aimed to fill gaps in the existing health care system, but it lacked a clear pathway structure and well-defined boundaries.

...a lot of people are starting to get frustrated because there’s just no clear pathway structure, anything of what we are and what we’re meant to be doing and where we can set those boundaries of that’s not our role. It just seems to be like take everything, deal with it when it’s there.

[Allied Health Professional]

All staff members were mindful that patients are eager to leave the hospital and might be willing to accept any form of home service. They were also aware of the hospital’s pressure to discharge patients promptly due to the scarcity of available beds.

I think so many of them are just so desperate to get out of hospital they’ll agree to anything. And I think that the hospital is so desperate to get them out of hospital as well that they will sign them up to whatever program will take them. [Leadership/Admin Team]

The wider context, especially in terms of the continuity of care, was discussed from service and interorganizational perspectives. Staff emphasized the importance of connecting with GPs before a patient’s discharge to ensure the continuity of care; however, this was rarely done in practice.

Especially with our frailty pathway, we’re meant to call the GP and talk about our frailty management plan once they’re discharged from our service, but that’s just been really difficult. Haven’t been able to get through to the GP. [Allied Health Professional]

Support and knowing how to connect patients with services following discharge were also discussed as a matter of importance.

...discharging patients from inpatient rehab, there’s only a couple of go-to places, but then it’s very different if patients are already out in the community. Because there’s a lot of options for further support that you can link them with. [Doctor]

Transferring patients back to the hospital when their health deteriorated presented challenges. Staff encountered resistance from patients who were reluctant to return to the hospital setting. Moreover, escalation pathways were impeded by overcrowded emergency departments and a shortage of available beds in the hospital. These issues were encountered by all clinical staff.

The difficulty has been once we identify they need to come in, what pathway they go towards. ED’S ramping, there’s no beds in [local hospital], and then you’ve got to take those measured risks. [Doctor]

Domain 7: Embedding and Adaptation Over Time
In general, all of the staff recognized the advantages that the service could provide to clinicians, patients, and the broader health care system (Multimedia Appendix 4). Nevertheless, they emphasized the need for planning that considered the distinctive and specialized requirements of virtual wards.

We try and introduce new models like technology and you’re trying to just do it on the edge of everything else that is happening but really it needs its own setup...that’s one of the challenges really, is trying to introduce a new model of care into an old system. [Leadership/Admin Team]

Staff were engaged with adapting and embedding the technology within the service in response to their patient’s needs. There was recognition that the service is in its early stages of development (Multimedia Appendix 4), and as a result, certain aspects are not yet finely tuned. While some staff regarded glitches as failures, many understood that creating a new service is a gradual process that demands refinement over time.

It’s changing that mindset, not seeing it as a failure. It’s just seeing it as a sidestep again for now...They see that as a failure, which it isn’t. Sometimes we’ve just got to fine tune things a little bit to be able to safely manage you at home. [Nurse]

Sensemaking and collective reflection were evident throughout the interviews, with strengths of the service being the implementation of Medtasker as a platform for team communication and accountability, and remote health monitoring to help clinicians escalate care. Adaptive action is ongoing to ensure that the technology is more dependable and the service is reaching its target patient cohort. The team is continually learning and identifying areas for adaption: for example, having a stock of common medications to trial patients on them, and having more technology applicable to an older cohort with frailty such as fall devices, echocardiograms, and telemetry.

Discussion
This paper addresses the knowledge gap related to the facilitators and barriers experienced by hospital staff during the implementation of a new VRW [2]. The service aimed to provide early discharge with a functional and medical care program for adults who are able to rehabilitate at home. The perspectives of staff are pivotal as previous research has suggested that clinician endorsement can account for a significant portion of the variability in the adoption, growth, and continuity of telehealth services [47-49].

The VRW team successfully managed complex patients in the community but highlighted how some conditions were more challenging than others, such as individuals with fluid retention or severe wounds. This was particularly pertinent for older
individuals, especially those with comorbidities and those who reside alone or in care homes, where they might lack the capacity to manage monitoring and communication equipment. Older people who were not proficient with digital technology were not disadvantaged in terms of quality of care or through digital exclusion [2], as the team was flexible in providing tailored care and would provide in-person services. Yet, it appears that for a virtual ward to achieve success, the service must possess a clear identity that is supported by its triage and intake process. A mismatch was described between staff expectations of medically stable patients focused on rehabilitation and the actual referrals of frail patients of varied case mix who were discharged early in response to hospital pressures. These findings underscore the importance of the consideration of patient characteristics prior to enrollment and the necessity of tailored care.

Regarding technology, interoperability was poor, a common issue in virtual care [50,51], and there were delays in pinpointing problems that impeded adoption. Nevertheless, some of these challenges appeared to be rooted in a lack of staff and patient training on correct technology use, an issue that appears common in new technology-supported health care models [15,16]. The brevity of the VRW program (typically 2 weeks) posed challenges, as there is limited time available for staff to adequately train patients, particularly older individuals who may need more support [52]. The substantial effort required may also seem disproportionate to the program’s short duration. The attitudes around technology were multifaceted. While there was a consensus that the technology was useful, particularly in the escalation of care, promotion of self-management, and team communication and task management, there was recognition that it sometimes created burden for the patient and carers and that it was not always accurate in detecting health decline [53-55]. In line with recent reports, simplifying hardware and software use were deemed crucial for success [2].

The primary value proposition encompassed offering patients the option to select their care location and increasing access to rehabilitation without waiting in hospital for a bed. This was particularly pertinent for older individuals, considering the potential risks associated with hospital admissions [10], as well as for patients residing far from rehabilitation hospitals. Opinions on risk were varied; yet, overall, there was agreement that the VRW did not entail greater risks but rather presented different risks compared to inpatient rehabilitation. A notable concern was that patients reestablished contact with their community care teams, including GPs, a factor that could make cohesive care challenging. Additionally, fewer chances for in-person physical assessment, monitoring, and therapy raised concerns about clinicians possibly missing early signs of decline or patients receiving less intensive treatment compared to inpatient wards.

In terms of adopters, changing staff culture toward virtual care was recognized as an area for improvement. Many staff still preferred face-to-face visits due to familiarity or ease compared to dealing with technology. Uncertainty surrounding workflows and professional roles raised concerns that clinical staff were being asked to take on technology or administrative tasks. Furthermore, there were changes in staffing to align with the complexity of the patient cohort including the employment of a pharmacist and more experienced nurses. There was also a recognition of a shift in care responsibility to the patient and their family. Staff particularly acknowledged the crucial role that caregivers play in the patients’ journey during virtual care. Little is known about carers roles in virtual models [2], and our ongoing work is exploring the experiences of patients and their family members who were referred to the VRW.

Clear pathways for early recognition of patient deterioration and appropriate escalation procedures were in place to ensure patient safety (eg, patients having a single phone number to call and the availability of remote monitoring equipment). However, challenges arose from wider pressures on public Australian hospitals, including overcrowded emergency departments and insufficient ward beds, which hindered patient transfers. Patients were reluctant to return to the hospital for these reasons. This holds significance because prior research has indicated that a major concern among staff regarding virtual care is the ability to swiftly admit patients to the hospital if their condition deteriorates [56]. Interorganizational collaboration, particularly handovers with GPs, occurred infrequently, leading to difficulties in seamless patient care transitions.

To our knowledge, this study addresses a significant gap in understanding by delving into the perspectives of hospital staff regarding the facilitators and obstacles in implementing a VRW. A key strength of this research lies in its qualitative methodology, allowing for a deep exploration of the topic within its natural context and enabling a nuanced examination of the multifaceted social, cultural, and environmental factors at play. Moreover, by centering the voices and experiences of hospital staff, an often-marginalized group in research as evidenced by the paucity of published literature, we aim to provide a platform for their narratives to be heard and valued.

Virtual wards represent a relatively novel concept, and there exists some ambiguity regarding their terminology [2]. Subsequent studies might prioritize the global standardization of the model, recognizing the importance of precise terminology in ensuring the generalizability of insights within the literature. Despite these variations, our research findings retain applicability to services that use monitoring and virtual technologies for postdischarge patient care.

To summarize, most staff acknowledged the benefits of the VRW service but emphasized the need for sufficient resource allocation and additional time to plan and implement the service and address early hurdles. In terms of health care access and equity, it appears that in a rehabilitation setting with a primarily older cohort, technology complements physical consultations but does not replace them and this must be factored into service design and delivery. We underscore that new health care models cannot exist on the fringes of traditional frameworks, as they require their own well-defined structure and setup with clear workflow and professional responsibilities.
Acknowledgments
This project is funded by the Australian Government 2022 Medical Research Future Fund Models of Care to Improve the Efficiency and Effectiveness of Acute Care Grant (2018573).

Conflicts of Interest
MC has a dual role with Flinders University and the South Adelaide Local Health Network.

Multimedia Appendix 1
Interview questions.
[DOCX File, 28 KB - aging_v7i1e54774_app1.docx ]

Multimedia Appendix 2
Additional analysis information and researcher credentials.
[DOCX File, 20 KB - aging_v7i1e54774_app2.docx ]

Multimedia Appendix 3
Coding tree.
[DOCX File, 16 KB - aging_v7i1e54774_app3.docx ]

Multimedia Appendix 4
Additional quotes from hospital staff interviews.
[DOCX File, 30 KB - aging_v7i1e54774_app4.docx ]

References


31. CC BY 4.0 deed. Creative Commons. URL: https://creativecommons.org/licenses/by/4.0/ [accessed 2024-06-19]


Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
GP: general practitioner
NASSS: Nonadoption, Abandonment, Scale-up, Spread, and Sustainability
NHS: National Health Service
PRISM: Practical, Robust Implementation and Sustainability Model
RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance
SALHN: Southern Adelaide Local Health Network
VRW: virtual rehabilitation ward
Phase Angle and Impedance Ratio as Indicators of Physical Function and Fear of Falling in Older Adult Women: Cross-Sectional Analysis

Danielle A Sterner, MS; Jeffrey R Stout, PhD; Kworweinski Lafontant, MS; Joon-Hyuk Park, PhD; David H Fukuda, PhD; Ladda Thiamwong, PhD

Corresponding Author: Kworweinski Lafontant, MS

Abstract

Background: Older adults experience a significant decline in muscle integrity and function with aging. Early detection of decreased muscle quality can pave the way for interventions to mitigate the progression of age-related physical declines. Phase angle (PhA) and impedance ratio (IR) are measures of muscle integrity, which can be assessed quickly via bioelectrical impedance analysis (BIA) and may be indicative of physical function.

Objective: This study aimed to characterize the relationships among handgrip strength (HGS), sit-to-stand (STS), BTrackS balance scores, fear of falling (evaluated using the Short Falls Efficacy Scale–International [Short FES-I]), and IR among community-dwelling older adult women classified as having a low or high PhA.

Methods: A cross-sectional analysis was conducted with 85 older women (mean age 75.0, SD 7.2 years; mean weight 71.0, SD 15.0 kg; mean height 162.6, SD 6.1 cm). To examine the influence of PhA on performance measures, participants were divided into 2 PhA groups: high (>4.1°; n=56) and low (≤4.1°; n=29). Data were nonnormative; hence, the Mann-Whitney U test was used to evaluate between-group differences, and Kendall τ coefficients were used to determine the partial correlations.

Results: The low PhA group had a significantly higher IR (mean 0.85, SD 0.03) than the high PhA group (mean 0.81, SD 0.03; r=92; P<.001). The high PhA group had superior HGS (mean 21.4, SD 6.2 kg; P=.007; r=0.36), BTrackS balance scores (mean 26.6, SD 9.5 cm; P=.03; r=0.30), and STS scores (mean 16.0, SD 5.5; P<.001; r=0.49) than the low PhA group (mean HGS 17.6, SD 4.7 kg; mean BTrackS balance score 37.1, SD 21.1 cm; mean STS score 10.7, SD 6.2). Both PhA and IR were significantly correlated with HGS and BTrackS balance, STS, and Short FES-I scores (P<.05). However, on adjusting for the whole sample’s age, only PhA was strongly correlated with HGS (τb=0.75; P=.003) and STS scores (τb=0.76; P=.002). Short FES-I scores were moderately correlated with IR (τb=0.46; P=.07) after controlling for age. No significant between-group differences were observed for height, weight, or BMI.

Conclusions: PhA and IR are associated with physical function and the fear of falling in older women. However, only PhA was significantly associated with physical function (HGS and STS) independent of age. Conversely, only IR was significantly associated with the fear of falling. Diminished physical function and increased IR appear to be characteristics of older women with a PhA of ≤4.1°. These findings suggest that PhA and IR measured through BIA together may serve as a valuable tool for early identification of older women at the risk of functional decline and a heightened fear of falling.

Trial Registration: ClinicalTrials.gov NCT06063187; https://clinicaltrials.gov/study/NCT06063187

International Registered Report Identifier (IRRID): RR2-10.2196/27381

KEYWORDS

handgrip; sit-to-stand; fitness; assessment; functionality; body composition; balance; fall; impedance; bioelectrical

Introduction

By 2060, the population of American adults older than 65 years is projected to surge, doubling from 52 million in 2018 to 95 million [1]. Muscle weakness and functional loss contribute to falls; in 2020, approximately 3 million adults older than 65 years were treated in emergency rooms as a result of falling, with 800,000 of them having been hospitalized [2]. The annual
cost of falls has exceeded US $50 billion, with Medicare and Medicaid covering 75% of these costs [2]. These older adults often experience a significant decline in muscle quality, quantity, and function [3]. Such age-related functional limitations can make everyday tasks, such as cooking, cleaning, and interacting with grandchildren, increasingly challenging [3]. The natural aging process brings about changes in body composition, often directly affecting physical function, as seen in sarcopenia [4]. Sarcopenia is characterized by a rapid loss of muscle strength with aging and is significantly associated with physical disability, decreased quality of life, and increased mortality [5,6]. Moreover, aging exacerbates traits of frailty, a syndrome characterized by a decrease in reserve capacity across various physiological systems, reducing their ability to withstand minor stressors [7,8]. This can result in an increased risk of falls, fractures, and disabilities [9,10], as well as higher mortality rates [7,8].

Aging also affects the musculoskeletal system, causing a decrease in skeletal muscle integrity, mass, strength, and function, along with an increase in the accumulation of noncontractile and adipose tissue [11,12]. Decreases in skeletal muscle integrity can result from a diminished cross-sectional area of muscle fibers, a transition from type II to type I fibers, and a loss of innervation [13]. A loss of skeletal muscle integrity can manifest through impaired balance and subsequently decreased physical function [13]. While not a physical characteristic, a fear of falling can also impose a restriction on physical activity and exacerbate functional loss [14]. Brouwer et al [15] assessed the fear of falling in healthy older adults and concluded that fear of falling was associated with poorer physical function characterized by a lower walking speed and lower limb weakness. Commonly assessed through questionnaires, early detection of the fear of falling can lead to interventions designed to increase physical activity in older adults and mitigate functional loss [16]. Likewise, early detection of decreased muscle quality can pave the way for interventions that may mitigate the clinical progression of sarcopenia and frailty [17].

Bioelectrical impedance analysis (BIA) is a quick, noninvasive, cost-effective method for assessing body composition within the Two-Compartment Model. It has also gained popularity, especially among older populations, as an assessment for skeletal muscle integrity and cellular health [10]. The BIA involves directing a constant low-level electrical current through the body [18]. This system has 3 main components: reactance (Xc), which measures the ability of cell membranes to store electrical charge (capacitance); resistance (R), which represents the resistive properties of cells due to intracellular water (ICW) and extracellular water (ECW) [19]; and impedance (Z), which represents the overall opposition to the electrical current [20]. Impedance provides another opportunity to examine cellular integrity with body cell mass [21]. Impedance ratio (IR) measures Z at high and low frequencies (in kHz) and can be indicative of possible cell membrane dysfunction based on body cell mass [22]. The IR for the whole body, upper limbs, and lower limbs is commonly calculated at 250 kHz or 5 kHz [21,23]. At higher frequencies, Z can penetrate cell membranes, therefore allowing for total body water (ECW + ICW) to be measured; however, at lower frequencies, Z can only measure ECW [24]. An IR ratio closer to 1 is indicative of cell membrane disruption, allowing more fluids, proteins, and electrolytes to shift into the extracellular space [22]. A strong inverse correlation has been reported between the phase angle (PhA) and IR in different clinical populations [22]. PhA is defined as the delay in current flow caused by a reduction in cell membrane capacitance [25].

PhA is calculated as the ratio of R to Xc, at a frequency of 50 kHz, as measured through BIA [26]. Furthermore, PhA is influenced by hydration status and lean body mass [27]. Therefore, it directly relates to the electrical functioning of cell membranes, skeletal muscle integrity, and PhA itself [12,27]. Higher PhA values are indicative of superior cell membrane integrity and cellular health [10]. Disease, dehydration, inflammation, malnutrition, and functional disabilities can cause disturbances in electrical tissue properties, reflected by a lower PhA [28-30]. A low PhA increases the risks for disability, falls, sarcopenia, frailty, and mortality among older populations [29]. PhA can be used as a proactive measure against physical weakening by identifying older individuals at the risk of muscle loss and mortality [10,17,19].

Previous research has established relationships between PhA and handgrip strength (HGS) [31], balance [13], gait speed [32], age [10], sex [30], sarcopenia [33], and BMI [34]. However, previously reported regression models have only been able to account for approximately ≤30% of the variance in PhA [31,35] and have largely left the fear of falling unaccounted for. Therefore, the primary objective of this study was to compare HGS, sit-to-stand (STS) scores, BTrackS balance scores, and fear of falling between groups of older women categorized as having high or low PhA. Additionally, we aimed to assess the relationships among PhA, IR, and physical function metrics while controlling for age, and determine which variables are most strongly associated with PhA and IR in this population.

We hypothesized that those with a high PhA would perform better on functional assessments and have a lower fear of falling, and that HGS would have the strongest association with PhA and IR when controlling for age.

**Methods**

**Ethical Considerations**

All study procedures were conducted in accordance with the tenets of the Declaration of Helsinki, approved by the University of Central Florida’s institutional review board (ID: STUDY00002473), and preregistered on ClinicalTrials.gov (NCT06063187).

**Recruitment**

The 85 female participants (n=64, 75% White; n=15, 17% Hispanic; n=4, 5% African American; n=2, 2% Asian) included in this analysis were part of a larger study funded by the National Institutes on Aging (RO3AG069799) [36]. This analysis used data from female participants only to specifically characterize the relationships between bioelectrical impedance parameters and physical function metrics in older women. This study used a cross-sectional design to determine whether there
were differences in HGS, STS, BTrackS balance scores, fear of falling, and IR based on a low or high PhA among older adult women. The sample of 85 older adult women had a mean age of 75.0 (SD 7.2) years, mean weight of 71.0 (SD 15.0) kg, and mean height of 162.6 (SD 6.1) cm. The study was conducted in low-income communities around central Florida, United States. Recruitment was achieved through flyer distribution, face-to-face engagement, local newsletters, and word of mouth. Community partners and clinical sites facilitated our introduction to potential participants for informed consent, initial screening, and eligibility verification using a checklist. Upon completion of the study, the participants received a US $30 gift card.

Eligible participants met all of the following inclusion criteria: (1) being aged ≥60 years; (2) being of low income status, based on 2019 poverty thresholds relative to family size and the number of children aged ≤18 years [18]; (3) absence of marked cognitive impairment, defined by a memory impairment screen score of ≥5 [37]; and (4) living independently in their own homes or apartments. Exclusion criteria were (1) medical conditions that inhibit balance testing or physical activity, such as the inability to stand on the balance plate or experiencing shortness of breath during physical activity; (2) active treatment from a rehabilitation facility; or (3) the presence of medical implants, such as pacemakers.

**Measurements**

**Grouping of Participants: Low and High PhA**

To verify the influence of PhA on performance measures, participants were divided into 2 groups: low PhA (≤4.1°; n=56) and high PhA (>4.1°; n=29). This cutoff was based on previous research that observed a higher prevalence of physical dysfunction and sarcopenia among community-dwelling women aged ≥65 years with a PhA less than 4.1° [34,38].

**BIA**

Body composition was assessed using the InBody s10, a direct segmental multifrequency BIA device from InBody Co, located in Seoul, South Korea. This device is designed to measure Z at 6 different frequencies—1, 5, 50, 250, 500, and 1000 kHz—for both the entire body and individual body segments. All BIA assessments were conducted before performing all other assessments, and all assessments were completed between 9 AM and 12 PM. To ensure accuracy, participants were instructed to fast for a duration of 3-4 hours, abstain from caffeine or alcohol for 24 hours, and avoid exercising for a period of 6-12 hours before testing. Participants were asked to maintain their normal dietary habits and arrive for testing adequately hydrated. On the day of assessment, participants were asked to remove their shoes, socks, and any metallic items. Height and weight were assessed using a digital physician scale and stadiometer (Health-O-Meter, Model 402 KL). Participants were then seated as their skin was prepared with an InBody wipe (InBody Co), and touch-type electrodes were then positioned on their left and right ankles, middle fingers, and thumbs. Participants were required to remain still for 1 minute before the electrodes were removed. PhA was derived by calculating the ratio of R to Xc at 50 kHz using the following formula: arc tangent (Xc/R) × (180/π) [26]. Moreover, IR was determined by dividing Z at 250 kHz by Z at 5 kHz [22,23]. The InBody s10 has good test-retest reliability among adults aged 65 years and older, with an intraclass correlation coefficient (ICC) of 0.82 [39].

**HGS**

Following BIA assessments, HGS was measured using a JAMAR Plus digital handgrip dynamometer (JLW Instruments) to ascertain maximal isometric force in kilograms. Participants, seated with feet flat on the floor and elbow bent at 90°, held the dynamometer in their hand, which was adjusted to allow for a flat second metacarpal and 90° bend at the knuckles. Participants then squeezed the dynamometer as hard as possible for 3-5 seconds across 3 trials, with 30-second rest intervals between each trial. All 3 trials were completed for 1 hand before performing 3 trials with the opposite hand. The maximum value for each hand was recorded, averaged, and used for analysis. The JAMAR handgrip dynamometer is a sound method to test HGS in the clinical setting with good intra- and interrater reliability [36,40].

**BTrackS Balance Scores**

BTrackS balance scores were gauged using the BTrackS balance system (Balance Tracking Systems) following HGS assessments. The BTrackS balance plate and BTrackS Assess Balance software (version 5.5.9) were used to measure center-of-pressure sway during a static stance. The scores and percentile rankings (0-100) were determined in accordance with age group and sex. A score of ≤30 signifies normal balance, while a score of ≥31 indicates poor balance and a moderate to high fall risk [41]. For each trial, participants were instructed to place their hands on their hips, close their eyes, and maintain a static position on the BTrackS Balance Plate for 20 seconds. Participants first underwent a familiarization trial that did not count toward their average score, followed by 3 trials that were averaged into their final score. To mitigate the risk of falls, a piece of sturdy furniture or a walker was placed within the participants’ reach during the test. The BTrackS balance system has been validated and has excellent reliability, with a Pearson correlation coefficient (r) of >0.90 and high test-retest reliability with an ICC of 0.83 [41].

**STS**

After BTrackS balance assessments, participants completed the 30-second STS test. STS scores were evaluated by instructing participants to stand up from a chair as many times as possible within 30 seconds. During the test, participants sat in the middle of the chair with their wrists crossed and hands resting on opposite shoulders. The 30-second STS test is a well-validated functional function measure in clinical research and practice, with good test-retest and interrater reliability [42].

**Short Falls Efficacy Scale–International**

Participants were asked to fill out the Short Falls Efficacy Scale–International (Short FES-I) questionnaire, which includes answering 7 questions on a scale of 1-4 that indicates if the participant would be concerned about falling during different activities. In the Short FES-I, a 1 indicates no concern at all and a 4 indicates being very concerned. The Short FES-I has been validated as a predictor of future falls and declines in functional
capacity with balance and gait, and it has excellent test-retest ability [16].

Statistical Analysis

All statistical analyses were conducted using SPSS (version 28; IBM Corp). Descriptive data are presented as mean (SD) values along with ranges where appropriate. The Mann-Whitney U test was used to evaluate between-group differences for all variables. A Kolmogorov-Smirnov test determined that the data were nonnormal, so Kendall τ coefficients were used to determine the partial correlations between variables controlled for age. The α value was set considering a P value of <.05.

Results

A total of 88 older women were screened for eligibility and 85 were included in the analysis, with 34% (29/85) of them in the low PhA group and 66% (56/85) of them in the high PhA group. As shown in Table 1, women in the low PhA group were significantly older than those in the high PhA group (P=.001). As anticipated, the group with a low PhA had a lower PhA (P<.001) and a higher IR (P<.001) than those in the high PhA group. In the high PhA group, PhA ranged from 4.2° to 7.0°, and IR ranged from 0.69 to 0.84. In the low PhA group, PhA ranged from 2.5° to 4.1°, and IR ranged from 0.83 to 0.89. No significant between-group differences in height, weight, or BMI were observed. Approximately 52% (44/85) of participants were considered overweight or obese with a BMI of ≥25.9 kg/m².

Table 1. Participants’ characteristics and bioelectric impedance analysis (n=85).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total (n=85), mean (SD)</th>
<th>Low phase angle (n=29), mean (SD)</th>
<th>High phase angle (n=56), mean (SD)</th>
<th>P valuea</th>
<th>Effect sizeb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>75.0 (7.2)</td>
<td>79.6 (8.2)</td>
<td>71.0 (5.6)</td>
<td>.001</td>
<td>0.43</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>162.6 (6.1)</td>
<td>162.0 (6.0)</td>
<td>163.0 (6.2)</td>
<td>.52</td>
<td>0.09</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>71.0 (15.0)</td>
<td>71.9 (15.7)</td>
<td>71.8 (14.8)</td>
<td>.81</td>
<td>0.03</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>26.8 (5.0)</td>
<td>27.3 (5.2)</td>
<td>26.9 (4.9)</td>
<td>.48</td>
<td>0.09</td>
</tr>
<tr>
<td>Body fat (%)</td>
<td>33 (10)</td>
<td>38 (9)</td>
<td>31 (10)</td>
<td>.003</td>
<td>0.40</td>
</tr>
<tr>
<td>Phase angle (°)</td>
<td>4.4 (0.8)</td>
<td>3.6 (0.4)</td>
<td>4.8 (0.6)</td>
<td>&lt;.001</td>
<td>1.00</td>
</tr>
<tr>
<td>Impedance ratio</td>
<td>0.82 (0.03)</td>
<td>0.85 (0.01)</td>
<td>0.81 (0.03)</td>
<td>&lt;.001</td>
<td>0.92</td>
</tr>
</tbody>
</table>

P values refer to the difference between the groups (Mann-Whitney U test).

Table 2 shows that women in the low PhA group demonstrated poorer physical function than those in the high PhA group. Specifically, the low PhA group had significantly lower average HGS (P=.007) and STS scores (P<.001). The low PhA group also showed significantly higher balance scores (P=.03) and Short FES-I scores (P=.001) than the high PhA group.

Table 2. Physical function parameters (n=85).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Low phase angle (n=29), mean (SD)</th>
<th>High phase angle (n=56), mean (SD)</th>
<th>P valuea</th>
<th>Effect sizeb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average handgrip strength</td>
<td>17.6 (4.7)</td>
<td>21.4 (6.2)</td>
<td>.007</td>
<td>0.36</td>
</tr>
<tr>
<td>(kg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sit-to-stand score</td>
<td>10.7 (6.2)</td>
<td>16.0 (5.5)</td>
<td>&lt;.001</td>
<td>0.49</td>
</tr>
<tr>
<td>BTrackS balance score</td>
<td>37.1 (21.1)</td>
<td>26.6 (9.5)</td>
<td>.03</td>
<td>0.30</td>
</tr>
<tr>
<td>Short FES-I score</td>
<td>11.6 (4.2)</td>
<td>9.3 (3.3)</td>
<td>.001</td>
<td>0.42</td>
</tr>
</tbody>
</table>

P values refer to the difference between the groups (Mann-Whitney U test).

Kendall rank correlation analysis (Table 3) revealed significant inverse correlations between PhA and age (τb=−0.37; P<.001) and between PhA and IR (τb=−0.79; P<.001). Significant moderate direct correlations were observed between PhA and STS scores (τb=0.34; P<.001) and between PhA and average HGS (τb=0.22; P=.002). Small but significant correlations were found between PhA and balance scores (τb=−0.19; P=.01) and between PhA and Short FES-I scores (τb=−0.25; P=.001). IR had a significant and direct relationship with age (τb=0.37; P<.001) and Short FES-I scores (τb=0.26; P<.001). IR had a significant and inverse relationship with average HGS (τb=−0.21; P=.004) and STS scores (τb=−0.33; P<.001).
Table. Relationships of phase angle and impedance ratio with participant characteristics (n=85).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Phase angle</th>
<th>Impedance ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau_b$</td>
<td>$P$ value $^b$</td>
</tr>
<tr>
<td>Age</td>
<td>-0.37</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Height</td>
<td>0.09</td>
<td>.25</td>
</tr>
<tr>
<td>Weight</td>
<td>0.06</td>
<td>.39</td>
</tr>
<tr>
<td>BMI</td>
<td>0.03</td>
<td>.64</td>
</tr>
<tr>
<td>Impedance ratio</td>
<td>-0.79</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Average handgrip strength</td>
<td>0.22</td>
<td>.02</td>
</tr>
<tr>
<td>Sit-to-stand score</td>
<td>0.34</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>BTrackS balance score</td>
<td>-0.19</td>
<td>.01</td>
</tr>
<tr>
<td>Short FES-I $^d$ score</td>
<td>-0.25</td>
<td>.001</td>
</tr>
</tbody>
</table>

$^a\tau_b$=Kendall $\tau_b$ correlation coefficient.  
$^bP$ values refer to the correlation between variables.  
$^c$Not available.  
$^d$Short FES-I: Short Falls Efficacy Scale–International.

After controlling for age (Table 4), strong direct correlations were observed between PhA and average HGS ($\tau_b=0.75; P=.003$) and between PhA and STS scores ($\tau_b=0.76; P=.002$). A large direct correlation was observed between IR and Short FES-I scores ($\tau_b=0.46; P=.07$). A moderate inverse correlation was found between IR and STS scores ($\tau_b=-0.32; P=.20$).

Table. Partial correlations between phase angle and impedance ratio (n=85).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Phase angle</th>
<th>Impedance ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau_b$</td>
<td>$P$ value $^b$</td>
</tr>
<tr>
<td>Height</td>
<td>0.50</td>
<td>.04</td>
</tr>
<tr>
<td>Weight</td>
<td>0.21</td>
<td>.41</td>
</tr>
<tr>
<td>BMI</td>
<td>0.10</td>
<td>.68</td>
</tr>
<tr>
<td>Average handgrip strength</td>
<td>0.75</td>
<td>.003</td>
</tr>
<tr>
<td>Sit-to-stand score</td>
<td>0.76</td>
<td>.002</td>
</tr>
<tr>
<td>BTrackS balance score</td>
<td>-0.04</td>
<td>.87</td>
</tr>
<tr>
<td>Short FES-F score</td>
<td>-0.24</td>
<td>.33</td>
</tr>
</tbody>
</table>

$^a\tau_b$=Kendall’s $\tau_b$ partial correlation coefficient.  
$^bP$ values refer to the correlation between variables, controlling for age.  
$^c$Short FES-I: Short Falls Efficacy Scale–International.

**Discussion**

**Principal Results**

The purpose of this study was to examine the relationships among HGS, STS score, balance, fear of falling, PhA, and IR in older adult women classified as having low or high PhA. When comparing physical function between high and low PhA groups, our results demonstrate significant differences in physical function between the high and low PhA groups. We observed a significantly lower IR within the high PhA group ($P<.001; r_{\tau_b}=0.92$) and a strong negative correlation between PhA and IR ($\tau_b=-0.79; P<.001$). The low PhA group had a significantly higher IR ($P<.001; r_{\tau_b}=0.92$) and lower PhA ($P<.001; r_{\tau_b}=1.0$).

When comparing physical function between the high and low PhA groups, our results demonstrate significant differences in physical function between groups. The low PhA group showed significantly lower HGS ($P=.007; r_{\tau_b}=0.36$), STS scores ($P<.001; r_{\tau_b}=0.49$), and higher balance scores ($P=.03; r_{\tau_b}=0.30$) than the high PhA group. Additionally, when examining relationships among PhA, IR, and physical function on controlling for age, moderate correlations were observed between PhA and HGS ($\tau_b=0.75; P=.003$) and STS scores ($\tau_b=0.76; P=.002$). However, balance scores ($\tau_b=-0.04; P=.87$) and fear of falling ($\tau_b=-0.24; P=.33$) showed only weak correlations with PhA when controlling for age. The low PhA group demonstrated a significantly higher Short FES-I score than the high PhA group ($P=.001; r_{\tau_b}=0.42$).
Comparison With Previous Literature

Previous literature has demonstrated an increased prevalence of physical dysfunction corresponding with a PhA less than $4.1^\circ$ in older community-dwelling women, which informed our cutoff value of $4.1^\circ$ to classify participants as having a low or high PhA [34,38]. Beyond a difference in PhA, the low PhA group also had a significantly lower IR of 0.85 (SD 0.01). This supports previous literature indicating that an IR closer to 1 is indicative of poor cellular health [22]. We observed a significantly lower IR in the high PhA group (Table 1) and a strong negative correlation between PhA and IR (Table 3). This aligns with previous evidence associating both a higher PhA and a lower IR with improved cellular integrity and health [22].

Reduced muscular strength and physical function in older adults has been associated with a lower PhA and higher IR [12,30]. In our study, we did not observe a strong correlation between HGS and PhA (Table 3). This is contrary to previous research, where IR and HGS were significantly correlated ($r=0.46$; $P<.001$) when controlling for age [24]. This discrepancy may be due to differences in methodology; Ballarin et al [24] assessed IR among 19-35-year-olds using a 300 kHz/5 kHz frequency ratio, while our study used a 250 kHz/5 kHz frequency ratio and included participants no younger than 60 years. IR is understood to be lower in younger populations [27], and younger populations still experience increases in HGS. This is contrasted by the higher IR and declining HGS experienced by older individuals.

Few studies have examined the relationship between STS score and PhA. Previous studies have instead used the gait speed test to assess physical function [43]. While STS and gait speed tests are not synonymous, both are dynamic multijoint movements that require both muscular strength and balance. This may explain why STS and gait speed performance are consistently observed as strong predictors of PhA in healthy individuals. This aligns with previous studies showing associations between lower limb strength and PhA [44]. Retaining muscle mass and physical function in the lower legs would have a direct impact on the R, Xc, and Z measured by BIA via an increase in muscular tissue and intracellular hydration [45]. A recent systematic review and meta-analysis of randomized controlled trials supports this theory, reporting that resistance training of at least 8 weeks increases PhA in older adults [45]. Furthermore, 6 out of the 7 studies included only involved female participants, which aligns with our sample [45]. Within clinical settings where the risks of conducting an STS test or other physical function assessments cannot take place, BIA may serve as a proxy for skeletal muscle quality and physical functioning. PhA has been shown to change with physical functioning longitudinally, as numerous studies have focused on sarcopenia and frailty regarding PhA [29].

The poorer physical function seen in the low PhA group suggests that a lower PhA reflects diminished skeletal muscle integrity and functionality in older adult women. Balance scores were significantly higher in the low PhA group, which are representative of poor balance and moderate to high fall risk (Table 1). A longitudinal study conducted by Asano et al [13] concluded that lower body strength diminishes with aging, and the observed poor balance score was associated with low PhA. We observed a greater fear of falling in the low PhA group (Table 2), and a moderate correlation between IR and fear of falling (Table 4). One study concluded that in women with osteoporosis, slower walking speed, decreased muscular strength, and greater postural sway were correlated with an increased fear of falling [46]. An increase in the fear of falling is associated with decreased muscular strength, which aligns with our findings [15].

Strengths and Limitations

One of the strengths of this study was the diverse population, including low-income female participants from 4 different racial and ethnic backgrounds. This study also intentionally used portable, accessible, and valid instruments to increase the applicability of the results to clinical practice. However, there are limitations to the study that should be considered. The main findings were raw BIA variables that are directly influenced by fluid distribution throughout the body, as the different frequencies used in calculated IR allow for the assessment of ECW and ICW. Hydration status may thus be a confounding variable as it was not assessed or controlled, although all testing occurred in the late morning for all participants after they were encouraged to void their bladders. In addition, the nonnormal data distribution was accounted for by using robust nonparametric tests during statistical analysis.

Implications and Future Directions

BIA is a brisk assessment that can be used in older adults to evaluate body composition and cellular health. This study concluded that PhA and IR are both linked to physical function and fear of falling in older women but associate differently when controlling for age. Although they are both measures of cellular health, our study demonstrates how PhA and IR differ in their relationship with physical function and fear of falling. As BIA continues to grow as a clinical assessment, there is a need to better understand how its measures relate to other assessments. A higher IR closer to 1 is indicative of poorer cellular health, which was observed in our low PhA group [22], and is also associated with diminished physical function and a heightened fear of falling in this study. Our study aligns with previous research reporting a strong inverse correlation between PhA and IR [21]; yet, there has been a lack of research investigating IR and its relationship with physical function. IR is a direct reflection of hydration status; therefore, we believe that IR is a helpful passive assessment to use, given the noninvasive and simple nature of the tool. More research on IR in older adults is needed to further examine its relationships with functional assessments.

Aging is associated with a change in body composition and decline in physical function. Therefore, PhA may reflect skeletal muscle health and can be assessed along with physical function. Based on this study, a PhA of $4.1^\circ$ would be indicative of decreased physical function in older women, and an intervention can be implemented to help improve PhA. Additionally, in instances where physical function assessments cannot take place, measuring PhA and IR may be valuable as indicators of physical function for that period. Our results suggest that muscular strength assessments such as HGS and STS may be more closely
related to PhA than balance in older adult women. STS scores are indicative of lower body strength, suggesting that lower body strength may be an important factor in PhA and IR. Balance scores can be indicative of muscle integrity, which is associated with frailty, sarcopenia, and malnutrition [29]. It is plausible that lower body strength may be particularly important when evaluating cellular health and physical function via bioimpedance parameters in older women, but more research is needed.

While physical function assessments are commonplace among older adults, psychological and physiological assessments should also be considered. Our results show that the fear of falling has an inverse relationship with physical function and PhA and a direct relationship with IR. Future research should further examine the relationships among fear of falling, physical function, PhA, and IR as they change over time. Understanding how these variables influence each other may aid in designing interventions to improve the health and quality of life of older adults.

**Conclusions**

Our results indicate that low PhA (≤4.1°) and high IR are linked with poorer physical function in older women, particularly for HGS and STS ability. PhA and IR are variables that can be assessed regularly during routine checkups and provide an indication of physical function and cellular health. Despite being indicative of cellular integrity and health, IR has not been widely studied in older adults. Assessing hydration status along with BIA measurements may help strengthen the design of future studies. Future research should also compare IR and physical function to our results and assess changes in IR longitudinally within older adults.

**Acknowledgments**

The research reported in this publication was supported by the National Institute on Aging (R03AG06799), the National Institute on Minority Health and Health Disparities (R01MD018025), and the Office of the Director, Chief Officer for Scientific Workforce Diversity (COSWD) Office (3R01MD018025-02S1) of the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The National Institutes of Health had no role in the design, execution, or analysis of the study and data herein.

**Data Availability**

The data that support the results of this study are available from the corresponding author upon reasonable request.

**Authors’ Contributions**

DAS, JRS, JP, and LT contributed to conceptualization. JRS, JP, and LT contributed to the methodology. DAS, JRS, JP, DHF, and LT contributed to the investigation. DAS, KL, DHF, and LT contributed to data curation. DAS, JRS, and KL contributed to formal analysis. DAS and KL contributed to writing (original draft preparation). DAS, JRS, KL, JP, DHF, and LT contributed to writing (review and editing). JRS, JP, DHF, and LT contributed to supervision. LT contributed to project administration and funding acquisition. All authors read and agreed to the version of the manuscript intended for publication.

**Conflicts of Interest**

KL received financial support from the National Institutes of Health (supplemental grant 3R01MD018025-02S1), as well as the McKnight Doctoral Fellowship from the Florida Education Fund. JRS, JP, and LT received financial support from the National Institute on Aging (R03AG06799), the National Institute on Minority Health and Health Disparities (R01MD018025), and the Office of the Director, Chief Officer for Scientific Workforce Diversity (COSWD) Office (3R01MD018025-02S1) of the National Institutes of Health. LT also received support from the National Science Foundation (NSF2222662). DAS and DHF have no conflicts of interest to report. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**References**


Abbreviations

BIA: bioelectrical impedance analysis
ECW: extracellular water
HGS: handgrip strength
ICC: intraclass correlation coefficient
ICW: intracellular water
IR: impedance ratio
PhA: phase angle
R: resistance
Short FES-I: Short Falls Efficacy Scale–International
STS: sit-to-stand
X: reactance
Z: impedance
Development and Usability Testing of a Mobile App–Based Clinical Decision Support System for Delirium: Randomized Crossover Trial

Jiamin Wang1,2,*, PhD; Meihua Ji2,*, PhD; Yuan Han3, MSN; Ying Wu2, PhD

*these authors contributed equally

Corresponding Author:
Ying Wu, PhD

Abstract

Background: The 3-Minute Diagnostic Interview for Confusion Assessment Method–Defined Delirium (3D-CAM) is an instrument specially developed for the assessment of delirium in general wards, with high reported sensitivity and specificity. However, the use of the 3D-CAM by bedside nurses in routine practice showed relatively poor usability, with multiple human errors during assessment.

Objective: This study aimed to develop a mobile app–based delirium assessment tool based on the 3D-CAM and evaluate its usability among older patients by bedside nurses.

Methods: The Delirium Assessment Tool With Decision Support Based on the 3D-CAM (3D-DST) was developed to address existing issues of the 3D-CAM and optimize the assessment process. Following a randomized crossover design, questionnaires were used to evaluate the usability of the 3D-DST among older adults by bedside nurses. Meanwhile, the performances of both the 3D-DST and the 3D-CAM paper version, including the assessment completion rate, time required for completing the assessment, and the number of human errors made by nurses during assessment, were recorded, and their differences were compared.

Results: The 3D-DST included 3 assessment modules, 9 evaluation interfaces, and 16 results interfaces, with built-in reminders to guide nurses in completing the delirium assessment. In the usability testing, a total of 432 delirium assessments (216 pairs) on 148 older adults were performed by 72 bedside nurses with the 3D-CAM paper version and the 3D-DST. Compared to the 3D-CAM paper version, the mean usability score was significantly higher when using the 3D-DST (4.35 vs 3.40; P < .001). The median scores of the 6 domains of the satisfactory evaluation questionnaire for nurses using the 3D-CAM paper version and the 3D-DST were above 2.83 and 4.33 points, respectively (P < .001). The average time for completing the assessment reduced by 2.1 minutes (4.4 vs 2.3 min; P < .001) when the 3D-DST was used.

Conclusions: This study demonstrated that the 3D-DST significantly improved the efficiency of delirium assessment and was considered highly acceptable by bedside nurses.

Trial Registration: Chinese Clinical Trial Registry, ChiCTR-IOR-17010368; https://www.chictr.org.cn/showproj.html?proj=17671

(JMIR Aging 2024;7:e51264) doi:10.2196/51264

KEYWORDS
delirium; 3D-CAM; older adults; clinical decision support system; nurse; 3-Minute Diagnostic Interview for Confusion Assessment Method-Defined Delirium

Introduction

Delirium is a common complication that is highly prevalent among hospitalized older adults; it can affect up to 40% of hospitalized older patients [1,2]. Delirium is associated with many adverse outcomes, including cognitive and functional impairment, increased hospital stay, care burden, and mortality, and therefore leads to increased care needs and poor prognosis [3,4]. The severity of the adverse consequences of delirium are positively correlated with the severity and duration of delirium [5]. Therefore, early recognition of delirium is essential for timely management of delirium to improve the patient’s prognosis [6].

Although many clinical guidelines and representative societies have recommended the use of standardized assessment tools for daily delirium screening in clinical practice, delirium is still
poorly recognized among hospitalized older adults [6,7]. A previous study has shown that up to 66% of older adults with delirium went unrecognized in routine practice [8]. The possible reasons for underdiagnosis of delirium include: communication barrier, inadequate use of the screening tools, and a lack of education on and conceptual understanding of delirium [9].

The Confusion Assessment Method (CAM) has been used as the reference standard for delirium screening [10]. However, the use of the CAM to assess delirium requires a combination of cognitive testing and subjective judgment based on clinical experience. Lemiere et al [11] found that the sensitivity on the daily use of the CAM by bedside nurses was only 30%. Therefore, to overcome the feasibility issues associated with the use of the CAM, Marcantonio et al [12] developed the 3-Minute Diagnostic Interview for CAM-Defined Delirium (3D-CAM) based on the key features of the CAM and simplified the evaluation process. The 3D-CAM was perceived as easy to use, and it had a sensitivity of 92% to 100% and a specificity of 88% to 94% among studies that were tested in different ethnic groups [12-15]. Kuczbarska et al [15] has identified that the 3D-CAM is appropriate for delirium assessment in general wards.

However, a previous study has shown that the use of the 3D-CAM by bedside nurses in routine practice demonstrated relatively poor usability, with multiple errors (such as human error, misunderstanding of item content, and incomplete or missing nursing records) during assessment [16]. The assessment of alteration on attention and thinking must be combined with cognitive assessment. During the process, nurses need to memorize, calculate, and make a judgment according to the patient’s response, which is prone to human errors. Meanwhile, the patient’s consciousness status needs to be compared with the assessment results at admission by reviewing the patient’s medical records. With these identified challenges, nurses often fail to complete the assessment successfully due to communication problems and difficulties in finding or loss of related materials. Moreover, due to the limited knowledge level on delirium among clinical nurses, adding prompts or cues to the items can effectively improve nurses’ identification of delirium, in addition to providing training on basic delirium knowledge before the use of the 3D-CAM [16]. Therefore, there is an urgent need to develop and adopt innovative ways to promote prompt delirium assessment in routine practice both effectively and accurately.

Several studies have reported that a clinical decision support system (CDSS) with some degree of autonomy may help solve this issue and improve the accuracy and adequacy of delirium assessment among bedside nurses [17,18]. Marcantonio et al [19] developed a brief app-based delirium identification tool, and it has shown good performance. Based on the Confusion Assessment Method for Intensive Care Unit (CAM-ICU), mobile apps have also been developed for delirium screening of patients in intensive care units (ICUs), and they showed acceptable usability and accuracy when used by bedside nurses [20-22]. Therefore, in this study, we aimed to develop the Delirium Assessment Tool With Decision Support Function Based on the 3D-CAM (3D-DST) and to evaluate its usability among older patients by bedside nurses.

### Methods

#### Ethical Considerations

The research protocol and secondary data analysis were approved by the institutional review committee of Capital Medical University (2015SY49). This was a substudy under a clinical trial, which was registered at the Chinese Clinical Registry (ChiCTR-IOR-17,010,368). Verbal informed consent was obtained from each participant before the start of the program, and participants had the right to withdraw during the study. The data were anonymized. The study protocol was safe and reliable and did not provide any compensation to the participants.

#### Design and Development of the 3D-DST

##### Previous Work

The 3D-CAM was translated into Chinese in a previous study and validated by nurse researchers; it showed acceptable sensitivity and specificity among hospitalized Chinese older patients [23].

##### Phase 1: Analysis of Problems in the Use of the 3D-CAM

**Paper Version**

The task walkthrough method was used to fully address the end users’ needs and achieve the overall goal of automatic delirium evaluation [16].

**Phase 2: Design of the 3D-DST**

In this phase, we first formed a multidisciplinary team that included experts with rich experience in delirium assessment. Bedside nurses, nursing researchers, software engineers, and user interface designers who were specialized in developing nursing information systems (NISs) were involved in the designing phase. The 3D-DST was designed and developed following the American Medical Informatics Association usability design principle [24]. Details of the design principle are shown in Multimedia Appendix 1.

#### Evaluation Process Analysis and Optimization of the 3D-DST

Based on the evaluation content and rules of the original 3D-CAM, this study scrutinized and analyzed the 3D-CAM evaluation process, identified the best path, and reoptimized and standardized the 3D-CAM evaluation process. By optimizing the 3D-CAM evaluation process, only the necessary paths to support the delirium evaluation were reserved, unnecessary links were eliminated, and the assessment was standardized to reduce intermediate errors; thus, a clinical decision-making system for delirium screening was established. Automatic evaluation logic jumps were incorporated into the system design according to the evaluation rules. Based on the problems identified by bedside nurses and the overall goals of the system, the key functional modules of the 3D-DST were initially drafted in mind-mapping software (Xmind software, version: 3.7.4.0; XMIND LTD).

https://aging.jmir.org/2024/1/e51264
User Interface Design of the 3D-DST

Overview

The interface design was completed in several steps. First, we used AxureRP (Axure Rapid Prototyping) to draft the logical diagrams, workflow, functional components, and user interfaces of the 3D-DST. Second, we made a web page to facilitate communication with engineers and team members. This was achieved via WuliHub (a domestic data hosting and sharing platform), and the interface prototype diagram and interaction components drawn by AxureRP were uploaded into this web browser–based demonstration scheme. Via WuliHub, a set of HTML files were generated based on the interface prototypes and interaction schemes drawn by AxureRP; they were compressed and uploaded to the platform for easy sharing. Third, following the American Medical Informatics Association interface design principle, the order of the evaluation content of each interface was determined based on the evaluation content and the optimal evaluation process of the 3D-DST. Finally, the prototype design of the evaluation interface of the 3D-DST was completed with the identified functions, including evaluator registration, log-in authentication, and user log out. Corresponding functional assessment modules were also developed, along with the result-reporting interface. The 3D-DST was designed by following the existing delirium assessment system [23], with easy access and a friendly display; for example, the patient’s identification was obtained by scanning their wristband, and different color selection, buttons, and information composition were used to increase its feasibility and usability.

Phase 3: Architecture and Development of the 3D-DST

In this phase, the system architecture, databases related to personal information and assessment data, and the user interfaces were identified. The 3D-DST was developed to fit on Android-compatible devices (Huawei nova 3), as the personal digital assistant used in clinical practice was largely based on the Android system rather than the iOS system in China. We used Java, Spring Boot, and RouYi-Vue to program the backend framework and Vue for the front-end framework. Mysql was used to formulate the databases. GitLab and Docker were used to release the 3D-DST. To make the 3D-DST system more stable, the model-view-viewmodel (MVVM) was used as the architecture scheme. The synchronization between view and model was completely automatic without human interference, the data maintenance was completely managed by the MVVM, and the operating environments were Linux and Windows [25,26]. To maintain information security, the 3D-DST set a cookie scheme with a time limit. The users needed to log in and verify their identity again when the cookie expired.

Usability Evaluation of the 3D-DST

Overview

The usability of the 3D-DST was evaluated using the acceptance questionnaire, and the results were compared to the 3D-CAM paper version when it was used by bedside nurses. The questionnaire included 6 domains regarding usability: perceived usefulness, ease of use, ease of learning, trustworthiness, intention to use, and satisfaction. The performance of the app was evaluated on 4 domains and compared to that of the 3D-CAM paper version, which included the successful completion rate of delirium assessment (proportion of nurses who completed the assessment correctly when they used the 3D-CAM paper version and the 3D-DST), evaluation completion time (time used to complete the assessment), the number of mistakes made during assessment against the results from a researcher, and satisfactory evaluation using an acceptance questionnaire.

Design, Setting, and Participants

This study was conducted among 72 bedside nurses from 3 tertiary hospitals in Beijing, China. Eligible patients from 3 internal medical wards (neurology, respiratory, and cardiology) and 1 surgical ward (orthopedic) of the study hospitals were assessed by participating bedside nurses using both the 3D-DST and the original 3D-CAM paper version. All participating patients and bedside nurses provided informed consent before study initiation. Following convenient sampling, bedside nurses who met the following criteria were included in the usability testing phase: (1) registered nurse with a valid license; (2) had more than 1 year of working experience and had been working continuously in their department for more than 3 months; and (3) willing to participate in this research. Hospitalized older adults who were aged 65 years or older and could communicate effectively in Mandarin were included in the study, and those with identified severe visual or hearing impairment were excluded from delirium assessment. Nurses who declined to participate during the study period were also excluded from the final analysis.

Usability Evaluation

The usability evaluation of both the 3D-CAM paper version and the 3D-DST was conducted using the usability testing questionnaire designed by Feng et al [27]. The content validity index scores for the 3 areas in the questionnaire (topic suitability, topic importance, and content clarity) were 1.00, 1.00 and 0.96, respectively. Since the dimensions for evaluating the usability of the CAM-ICU and 3D-CAM are similar, the questionnaire can also be used to evaluate the usability of the 3D-CAM. In this study, we revised several items of the questionnaire to make it suitable for the 3D-CAM paper version or the 3D-DST. For example, we changed the item “I think this tool meets the requirements of ICU nurses for a delirium assessment tool” to “I think this tool meets the requirements of nurses for a delirium assessment tool in general wards.” In addition, the term “CAM-ICU” in the items evaluating the ease of use was replaced with “3D-CAM.” The reliability of the revised questionnaire was tested for overall internal consistency, and the Cronbach α coefficient was 0.907. The questionnaire was rated on a 5-point Likert scale ranging from “1=strongly disagree” to “5=strongly agree.” A higher score indicated better user acceptance. The usability evaluation was evaluated on 6 domains: perceived usefulness, ease of use, ease of learning, trustworthiness, intention to use, and user satisfaction. To evaluate the acceptability of the 3D-DST, items to evaluate the interfaces of the app were added to the questionnaire, so the
The usability questionnaire contained 26 and 43 items for the 3D-CAM paper version and the 3D-DST, respectively.

Before study initiation, researchers used a computer program to generate random numbers and made an assignment sequential table. Eligible bedside nurses were numbered according to the order they participated in. One group of nurses (Group A) used the 3D-CAM paper version first to evaluate 3 patients (initial evaluation at admission and follow-up assessments during the patients’ hospitalization, including assessments of possible cognitive impairment or identified delirium). Subsequently, the same group of bedside nurses used the 3D-DST to evaluate patients with a similar admission diagnosis after 24 hours to avoid the impact of short-term memorization on the evaluation process. Conversely, nurses from the other group (Group B) used the 3D-DST first and then the 3D-CAM paper version second to evaluate 2 different sets of patients (3 patients per nurse). Nurses who participated in the study only assessed patients admitted in the wards where they worked. A nurse researcher conducted training sessions for all bedside nurses on the use of the 3D-CAM and the 3D-DST before their assessments; the nurse researcher also selected the eligible patients based on their admission diagnosis prior to being approached. Usability testing of both the 3D-CAM paper version and the 3D-DST was carried out when each set of assessment was completed (Figure 1).

**Figure 1.** Flowchart of participant recruitment and participation. 3D-CAM: 3-Minute Diagnostic Interview for Confusion Assessment Method–Defined Delirium; 3D-DST: Delirium Assessment Tool With Decision Support Function Based on the 3-Minute Diagnostic Interview for Confusion Assessment Method–Defined Delirium.
Performance of the 3D-DST

During each assessment, 1 experienced nurse researcher observed the performance of bedside nurses and completed the patients’ delirium screening using the 3D-CAM paper version at the same time. The successful completion rate of delirium assessments in terms of the proportion of correctly identified delirium, whether using the 3D-CAM paper version or the 3D-DST by bedside nurses, was recorded. In addition, the number of mistakes made during the assessments (compared to the researcher’s assessment results) and the evaluation completion time were also recorded for each nurse. Nurses were blinded to the researcher’s assessment results.

Sample Size

The highest score for each item of the usability questionnaire was 5 points (the higher the score, the better the usability). We expected that an average score of 4 or more would be achieved when the 3D-DST was used by bedside nurses, which is 1 point higher than the average score of the 3D-CAM paper version. With the SD being 1.0 and α=.05, at least 54 bedside nurses were needed to achieve 90% power. Considering that a quarter of bedside nurses might not be able to complete the study, a final sample of 72 nurses was required for this study.

Data Analysis

SPSS software (version 21.0; IBM Corp) was used to perform the data analysis. Normally distributed variables were presented using the mean and SD, whereas nonnormally distributed variables were presented using the median and range. Categorical variables were presented with frequencies and proportions as appropriate. The χ² test or Fisher exact test was used to compare the differences between nurses’ baseline data (different age groups, sex, etc) and the evaluation success rate. The comparison of evaluation completion time, the total usability score, and the impact of the sequential use of the 3D-CAM paper version and the 3D-DST between 2 groups were achieved using ANOVA. A nonparametric test was used to test the differences between the scores of each domain of the usability questionnaire and the order of using the 2 types of assessments. P<.05 was considered statistically significant.

Results

Design and Development of the 3D-DST

Phase 1: Analysis of Problems in the Use of the 3D-CAM Paper Version

Several problems were identified with the use of the 3D-CAM paper version by bedside nurses, including human errors, insufficient or incorrect understanding of the assessment contents, and incomplete or failed retrieval of the relevant information from the nursing record [16]. In developing the 3D-DST, the delirium evaluation process was simplified as unnecessary steps were automatically omitted after sorting out the procedural results related to specific features of the 3D-CAM. The evaluation processes of the 3D-CAM paper version and the 3D-DST are shown in Multimedia Appendices 2 and 3.

Phase 2: Design of the 3D-DST

The 3D-DST was installed on mobile phones with an Android (8.1.0) system, 128 GB, 8-core processor, and 1.8 GHz. Three evaluation modules were incorporated into the 3D-DST, reflecting the inquiry, observation, and selective evaluations. The evaluation interfaces (Figure 2) of the 3D-DST included 8 evaluation pages and 16 different evaluation result interfaces (3 delirium-positive interfaces and 13 delirium-negative interfaces).
Auxiliary reminders based on the common mistakes reported by nurses in phase 1 were added to the system. In overcoming these burdens, the 3D-DST was developed to include reminders on the correct answers regarding items 1, 2, and 5. Nurses only needed to check whether the patient’s response was consistent with the reminder or not. As the 3D-DST can automatically record and retrieve previous evaluation results, item 22 was not displayed in the 3D-DST but was incorporated into the algorithm of the 3D-DST, and the result was generated automatically. To address issues associated with “incorrect understanding of item content,” cues were incorporated into the 3D-DST to facilitate better understanding of the contents of related items (11 through 20) during assessment.

In the development of the 3D-DST, we also included functions to automatically retrieve necessary information from the record for comparison, which was incorporated to the 3D-DST. Nurses using the 3D-DST do not need to manually search and compare the previous evaluation results, which could increase the successful completion rate of the assessment.

**Usability Evaluation of the 3D-DST**

A total of 72 bedside nurses completed the usability testing of the 3D-DST. The demographic characteristics of participating bedside nurses are shown in [Table 1](#). They were mainly female (n=67, 93%), with 44 (61%) nurses holding a bachelor’s degree. A total of 148 patients were evaluated by nurses in this study,
with 98 (66%) being female and a mean age of 76 (SD 7.18) years.

Table. Demographic characteristics of the participating bedside nurses.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Group A (n=36), n (%)</th>
<th>Group B (n=36), n (%)</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age group (years)</strong></td>
<td></td>
<td></td>
<td>.93</td>
</tr>
<tr>
<td>20-30</td>
<td>20 (56)</td>
<td>20 (56)</td>
<td></td>
</tr>
<tr>
<td>31-47</td>
<td>16 (44)</td>
<td>16 (44)</td>
<td></td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td>&gt;.99</td>
</tr>
<tr>
<td>Female</td>
<td>33 (92)</td>
<td>34 (9)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3 (8)</td>
<td>2 (56)</td>
<td></td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
<td></td>
<td>&gt;.99</td>
</tr>
<tr>
<td>Junior college</td>
<td>12 (33)</td>
<td>13 (36)</td>
<td></td>
</tr>
<tr>
<td>College and above</td>
<td>24 (67)</td>
<td>23 (64)</td>
<td></td>
</tr>
<tr>
<td><strong>Work experience (years)</strong></td>
<td></td>
<td></td>
<td>.71</td>
</tr>
<tr>
<td>&lt;5</td>
<td>15 (42)</td>
<td>16 (44)</td>
<td></td>
</tr>
<tr>
<td>5-10</td>
<td>11 (31)</td>
<td>9 (25)</td>
<td></td>
</tr>
<tr>
<td>( \geq 11 )</td>
<td>10 (28)</td>
<td>11 (31)</td>
<td></td>
</tr>
</tbody>
</table>

A total of 432 delirium assessments (216 pairs) were performed by the bedside nurses on 148 older adults with the 3D-CAM paper version and the 3D-DST, of which 38 (26%) older adults were confirmed as delirium positive by a nurse researcher using the 3D-CAM paper version.

The mean usability scores of bedside nurses using the 3D-CAM paper version and the 3D-DST were 3.40 (SD 0.43) and 4.35 (SD 0.31), respectively, with the score of the 3D-DST being 0.95 points higher than that of the 3D-CAM paper version (\( P<.001 \)). The median scores of the 6 domains of the satisfactory evaluation questionnaire for nurses using the 3D-CAM paper version and the 3D-DST were above 2.83 points and 4.33 point, respectively. As shown in Table 2, the median scores on the ease of use, ease of learning, and trustworthiness were 1 point higher than the other 3 domains (\( P<.001 \)). The total satisfactory score of the 3D-DST was higher than that of the 3D-CAM paper version (\( P<.001 \)), whereas the order of assessments, whether using the 3D-CAM paper version first or the 3D-DST first, had no effect on the results.
The assessment success rate of the 3D-CAM paper version was a little lower compared to that of the 3D-DST (203/216, 94% vs 212/216, 98.1%; \( P = .045 \)). The median time of assessment using the 3D-CAM paper version was 2.1 minutes longer than that of the 3D-DST (4.4 vs 2.3 min; \( P < .001 \)). The overall performances of the 3D-DST and the 3D-CAM paper version are displayed in Table 3.

### Table 3. Comparison of the effectiveness and efficiency of the 3D-CAM\(^a\) paper version and the 3D-DST\(^b\).

<table>
<thead>
<tr>
<th>Domain</th>
<th>Score, median (range)</th>
<th>Z value</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perceived usefulness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D-DST</td>
<td>4.57 (3.29-5.00)</td>
<td></td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td>3D-CAM paper version</td>
<td>3.86 (2.71-4.86)</td>
<td>-6.903</td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td><strong>Perceived ease of use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D-DST</td>
<td>4.33 (3.00-5.00)</td>
<td></td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td>3D-CAM paper version</td>
<td>2.83 (1.83-4.33)</td>
<td>-6.973</td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td><strong>Learnability</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D-DST</td>
<td>4.33 (3.00-5.00)</td>
<td></td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td>3D-CAM paper version</td>
<td>3.33 (2.33-4.67)</td>
<td>-7.014</td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td><strong>Trustworthiness</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D-DST</td>
<td>4.33 (3.00-5.00)</td>
<td></td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td>3D-CAM paper version</td>
<td>3.33 (2.00-5.00)</td>
<td>-6.697</td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td><strong>Intention to use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D-DST</td>
<td>4.40 (3.20-5.00)</td>
<td></td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td>3D-CAM paper version</td>
<td>3.67 (1.33-5.00)</td>
<td>-6.446</td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td><strong>Satisfaction</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3D-DST</td>
<td>4.33 (3.33-5.00)</td>
<td></td>
<td>( P &lt; .001 )</td>
</tr>
<tr>
<td>3D-CAM paper version</td>
<td>3.60 (1.60-4.20)</td>
<td>-6.985</td>
<td>( P &lt; .001 )</td>
</tr>
</tbody>
</table>

\( ^a \)3D-CAM: 3-Minute Diagnostic Interview for Confusion Assessment Method–Defined Delirium.

\( ^b \)3D-DST: Delirium Assessment Tool With Decision Support Function based on the 3-Minute Diagnostic Interview for Confusion Assessment Method–Defined Delirium.

### Table 4. Comparison of the performances between the 3D-CAM\(^a\) paper version and the 3D-DST\(^b\) by bedside nurses.

<table>
<thead>
<tr>
<th>Performance</th>
<th>Nurses, n</th>
<th>Assessments, n</th>
<th>Group A</th>
<th>Group B</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Successful completion rate, n/N (%)</strong></td>
<td>72</td>
<td>216</td>
<td>103/108 (95.4)</td>
<td>100/108 (92.6)</td>
<td>.045</td>
</tr>
<tr>
<td>3D-CAM paper version</td>
<td>72</td>
<td>216</td>
<td>107/108 (99.1)</td>
<td>105/108 (97.2)</td>
<td></td>
</tr>
<tr>
<td>3D-DST</td>
<td>72</td>
<td>216</td>
<td>107/108 (99.1)</td>
<td>105/108 (97.2)</td>
<td></td>
</tr>
<tr>
<td><strong>Human errors, n</strong></td>
<td>72</td>
<td>N/A(^c)</td>
<td>6</td>
<td>9</td>
<td>.62</td>
</tr>
<tr>
<td>3D-CAM paper version</td>
<td>72</td>
<td>N/A(^c)</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>3D-DST</td>
<td>72</td>
<td>N/A(^c)</td>
<td>2</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td><strong>Evaluation completion time (min), median (IQR)</strong></td>
<td>72</td>
<td>203</td>
<td>4.45 (2.5-5.4)</td>
<td>4.35 (2.4-5.3)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>3D-CAM paper version</td>
<td>72</td>
<td>212</td>
<td>2.25 (1.25-3.55)</td>
<td>2.35 (1.4-3.6)</td>
<td></td>
</tr>
<tr>
<td>3D-DST</td>
<td>72</td>
<td>212</td>
<td>2.25 (1.25-3.55)</td>
<td>2.35 (1.4-3.6)</td>
<td></td>
</tr>
</tbody>
</table>

\( ^a \)3D-CAM: 3-Minute Diagnostic Interview for Confusion Assessment Method–Defined Delirium.

\( ^b \)3D-DST: Delirium Assessment Tool With Decision Support Function Based on the 3-Minute Diagnostic Interview for Confusion Assessment Method–Defined Delirium.

\( ^c \)N/A: not applicable.
Discussion

Principal Findings

This study described the development process of a CDSS based on the 3D-CAM and evaluated its usability in delirium screening among older patients. Our results demonstrated that the 3D-DST was perceived as highly satisfactory with acceptable usability when used by bedside nurses, and it improved the completion rate and reduced the evaluation completion time when bedside nurses used the app among older patients.

CDSSs are tools incorporated with a significant clinical knowledge base and are designed to provide users (health care professionals, patients, and caregivers) with an intelligent way to assist in clinical decision-making [28]. Previous studies have found that well-designed CDSSs are effective and can improve clinical outcomes and health processes [29]. Therefore, it is very important to ensure the quality of CDSSs and avoid unpleasant situations when deploying unreliable systems.

Functional suitability is a very important feature when developing CDSSs; it refers to the extent to which a system meets the stated and implied requirements through its functional components under certain conditions [30-32]. In our study, the 3D-DST is aimed to connect with the hospital information system or NIS, so that it can automatically retrieve the patients’ information via the hospital information system. The 3D-DST can obtain information by scanning the QR code on the patient’s wristband, which allows the 3D-DST to be easily integrated into the portable NIS and facilitates efficient delirium assessment by bedside nurses.

As indicated by the study results, the median scores on the ease of use, ease of learning, and trustworthiness of the 3D-DST were over 1 point higher than those of the 3D-CAM paper version. It was demonstrated that the development of the 3D-DST met the requirements of bedside nurses, improving the acceptability and usefulness of the screening tool. There were several possible reasons. First, the 3D-DST was designed through process optimization, which incorporated strategies such as automatic evaluation logic jumps, embedded prompts, automatic comparison function, etc. The design process made full use of information technologies, such as automatic recording, calculation, and other intelligent functions, that could effectively solve the problems identified by nurses when they used the 3D-CAM paper version. The 3D-DST was well accepted by bedside nurses, the burden on the memorization of information was reduced, and the ease of use of the 3D-CAM was improved. Second, nurses only needed to complete the evaluation process by following the interfaces and the embedded prompts. Nurses did not need to learn the specific instructions, and the system could automatically record, calculate, and output the results, therefore improving this tool’s learnability and scalability. Third, since the content of the 3D-CAM is mostly a routine assessment, nurses needed less training time to use the tool. Moreover, with reduced time to complete the delirium assessment, nurses perceived the 3D-DST as highly acceptable, which generated trustworthiness and solved the existing problems associated with the use of the 3D-CAM paper version.

The improvement of the perceived usefulness, intention to use, and satisfaction scores was less than 1 point when comparing the 3D-DST with the 3D-CAM paper version. This could be attributed to the following reasons. The scores on these 3 domains of the 3D-CAM paper version were considerably high. Perceived usefulness was mainly evaluated based on nurses’ knowledge of delirium assessment, whereas intention to use mainly referred to whether nurses were willing to use the tool and whether the nurses could accept the method of using the tool or not, notwithstanding the evaluation completion time. Since this study was only conducted in a short period of time, nurses may not fully understand and appreciate that the 3D-DST can standardize the assessment process and improve the recognition rate and accuracy of the delirium assessment. In the future, the duration that the nurses use the 3D-DST should be extended before the acceptance evaluation. Furthermore, the intention to use the 3D-DST and satisfaction perceived by nurses may be affected by organizational factors.

Our study found that the 3D-DST had a slightly higher successful completion rate than the 3D-CAM paper version in assessing delirium among bedside nurses (212/216, 98.1% vs 203/216, 94%; P=.045). In the 3D-DST, reminders of the correct responses for items 1, 2 and 5 (objectively testing patients’ cognition) were incorporated into the app, which can increase the accuracy of the delirium assessment, reduce the information processing time, and improve nurses’ work efficiency. When nurses evaluate delirium using the 3D-DST, the system will automatically record and generate the evaluation results without manual input; this prevented possible human errors that are introduced by the nurses. The 3D-DST also added prompts that aimed to reduce the assessment failure rate caused by incorrect or inadequate understanding of the content of the items. By integrating prompts into the 3D-DST, it may have contributed to the reduced failure rate among bedside nurses by guiding and standardizing the assessment process; thus, insufficient knowledge and possible human errors are fully addressed during the assessment. Therefore, the 3D-DST improved the success rate of the delirium assessment. With assistance of CDSSs, decision-making can be incorporated into the routine assessment to guide nurses to complete the assessment successfully with standardized procedures [33-35].

The advantages of the study are multifactorial. The 3D-DST was developed to solve the problems associated with the use of 3D-CAM paper version in routine practice by nurses. The research team included multidisciplinary members such as delirium assessment experts and software development engineers, and we used a combination of multiple architectures to ensure the stability of the system during the development process. In addition, in evaluating the usability and performance (completion rate, time required for completing the assessment, and the number of human errors made by nurses during the assessment) of the 3D-DST, both subjective and objective approaches were used by comparing the 3D-CAM paper version with the 3D-DST among bedside nurses, which showed the promising results of the 3D-DST.

This study also has several limitations. First, each bedside nurse only used the 3D-CAM to evaluate 6 patients, which may have limited the nurses’ possibility to evaluate all types of patients.
with different admission diagnoses in general wards. Therefore, our result may not be generalizable to other clinical settings. Second, this study only applied a quantitative approach to evaluate the usability of the 3D-DST; interviews can be added to explore the usability of the 3D-DST among bedside nurses in the future. Third, during the usability testing phase, some usability issues in our system may not be adequately reflected due to limitations in the patients’ admission types and the limited number of assessments. Fourth, bedside nurses were not physically involved in the interface design process, which may have affected the usability of the interface. However, we have examined the problems and issues associated with the use of the 3D-CAM paper version among bedside nurses, and the 3D-DST was developed to address these problems by including experienced nursing researchers during the development process, so the results should not be affected. Finally, this study did not analyze the accuracy of the 3D-DST in assessing delirium, as this was not required for the study objectives at this stage, and it was completed as a separate study.

Conclusion
This study demonstrated that the 3D-DST was perceived as highly acceptable and useful in assisting bedside nurses to identify delirium among older adults in routine practice. The integration of this app with existing health systems could enhance its positive impact on the efficiency and accuracy of delirium screening in the future.

Acknowledgments
This study was supported by the National Natural Science Foundation of China (NSFC; grant 72034005). We would like to thank all the patients and nurses who participated in this study.

Authors’ Contributions
JW contributed to conceptualization, methodology, software, and writing—original draft. MJ contributed to writing—review and editing. YH contributed to conceptualization, methodology, software, validation, formal analysis, and investigation. YW contributed to conceptualization, methodology, software, validation, resources, writing—review and editing, supervision, project administration, and funding acquisition.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Details of the design principle.
[DOCX File, 13 KB - aging_v7i1e51264_app1.docx ]

Multimedia Appendix 2
Evaluation process of the original 3-Minute Diagnostic Interview for Confusion Assessment Method–Defined Delirium (3D-CAM) paper version.
[PNG File, 162 KB - aging_v7i1e51264_app2.png ]

Multimedia Appendix 3
[PNG File, 330 KB - aging_v7i1e51264_app3.png ]

Checklist 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File, 1742 KB - aging_v7i1e51264_app4.pdf ]

References


https://aging.jmir.org/2024/1/e51264


Abbreviations

3D-CAM: 3-Minute Diagnostic Interview for Confusion Assessment Method–Defined Delirium

3D-DST: Delirium Assessment Tool With Decision Support Function Based on the 3-Minute Diagnostic Interview for Confusion Assessment Method–Defined Delirium

CAM: Confusion Assessment Method

CAM-ICU: Confusion Assessment Method for Intensive Care Unit

CDSS: clinical decision support system

ICU: intensive care unit

MVVM: model-view-viewmodel

NIS: nursing information system

© Jiamin Wan, Meihua Ji, Yuan Han, Ying Wu. Originally published in JMIR Aging (https://aging.jmir.org), 24.1.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Digital Storytelling Intervention for Enhancing the Social Participation of People With Mild Cognitive Impairment: Co-Design and Usability Study

Di Zhu1,2, MAP; Abdullah Al Mahmud1*, PhD; Wei Liu2*, PhD

1Centre for Design Innovation, Swinburne University of Technology, Melbourne, Australia
2Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Centre for Experimental Psychology Education, Beijing Normal University, Faculty of Psychology, Beijing Normal University, Beijing, China

* these authors contributed equally

Abstract

Background: Community-based social participation has shown promise in delaying cognitive decline in older adults with mild cognitive impairment (MCI) who are at risk of developing dementia. Although group storytelling interventions have proven effective, the need for a skilled workforce to support people with MCI can limit broader community implementation. Technology-based interventions may offer a solution to this limitation by replicating the abilities of therapists.

Objective: This study aims to co-design a digital storytelling intervention and evaluate its usability.

Methods: This co-design process involved 3 stages, engaging people with MCI (n=12), their caregivers (n=4), and therapists (n=5) in Beijing, China. In the first stage, we used card sorting and voting methods to identify potential incentives for social participation and target the specific abilities that people with MCI wanted to enhance. In the second stage, we conducted brainstorming sessions with people with MCI and their caregivers to identify the potential features of a digital storytelling application named Huiyou (“meeting new friends” in Chinese). Finally, we assessed Huiyou’s usability with people with MCI and therapists, leading to iterative improvements based on the usability findings.

Results: We uncovered a crucial link between boosting the self-confidence of people with MCI and their ability to address social participation challenges. Notably, we identified memory improvement and enhanced language expression as key factors for effective communication with grandchildren. Subsequently, participants suggested features and interfaces to address these challenges, leading to the development of Huiyou, a group-based digital storytelling application featuring functions such as generating story materials, conducting memory retrieval activities, and sharing stories. It received an “excellent” rating in the User Experience Questionnaire benchmark, displaying high levels of attractiveness, dependability, stimulation, and novelty. People with MCI achieved an average task completion rate of 87% (n=19; SD 0.13) of the 22 tasks. However, feedback from people with MCI and therapists highlighted usability issues in navigation, activity management, user interface, and feature optimization, indicating a need for improved accessibility and efficiency.

Conclusions: The co-design approach contributed to developing the Huiyou prototype, supporting community-based social participation. User feedback highlighted the potential of Huiyou to enhance well-being and facilitate meaningful social interactions while maintaining crucial existing relationships.

(JMIR Aging 2024;7:e54138) doi:10.2196/54138

https://aging.jmir.org/2024/7/e54138
KEYWORDS
codeign; digital storytelling; people with mild cognitive impairment; MCI; technology-based intervention development; dementia; mobile phone

Introduction

Background
Social participation has proven efficacious in decelerating the progression of dementia from mild cognitive impairment (MCI) [1], particularly in low- and middle-income countries [2]. To enhance social participation among older adults with MCI, researchers are engaged in reducing the obstructions to social participation [3] and creating and customizing a variety of constructive social activities [4]. To expand social participation, researchers have explored an assortment of interventions aimed at improving memory, communication, and familial relationships among older adults with MCI [5]. Social participation is considered at different levels: societal, community, relationship, and individual levels [6]. Community involvement is defined by its significance and the sense of satisfaction it generates in people’s lives as well as its potential for creating social support [7]. Older adults reported that technology could assist them in maintaining social connections [8]. For example, researchers used cameras, enabling people with MCI to memorialize their daily lives [9]. Most interventions concentrate on enhancing cognitive abilities [10] or facilitating recreational and sports activities [11]. The former relies upon participants’ strong motivation to participate in interventions, whereas the latter has less impact on cognitive ability. Storytelling tools can simultaneously enhance cognitive abilities and create a novel form of social interaction. Incorporating storytelling as a therapeutic intervention possesses enormous potential for improving health care outcomes and promoting well-being among people with MCI [12]. Studies have demonstrated that socially isolated people with MCI may experience lower psychological well-being and exhibit more negative states of mind than older adults considered to have cognitive normality [13]. Owing to cognitive impairments, people with MCI may encounter the following issues: (1) reduced interest in and initiation of social interactions, suggesting a lack of proactive engagement [14]; (2) difficulties in establishing and maintaining interpersonal relationships, implying an impaired sense of social synchrony; and (3) challenges in recognizing and adhering to social boundaries and norms [15]. It is imperative to explore behavior change strategies aimed at fostering a positive social health lifestyle among people with MCI. Technology could be a beneficial tool to maintain social connection [8]. For example, ElderConnect is a web-based app designed to assist senior citizens in recognizing, preventing, and easing feelings of loneliness. It offers information and tactics to help them establish new social connections and sustain existing ones [16]. Notably, storytelling serves as an essential activity for promoting positive social health [17,18]. Harnessing the power of storytelling, people with MCI can be encouraged to participate more proactively in their health care by identifying their unique needs and knowledge gaps while fostering strong connections and support networks with peers with similar disease-related experiences [19].

However, these storytelling interventions often lack a focus on community-based social activities, primarily being individual interventions without emphasis on interpersonal interactions. Furthermore, community-based programs depend on skilled facilitators, such as therapists or social workers, to maintain quality. For instance, providing training for facilitators before they lead a group reminiscence program ensures expertise in techniques and effective program management [20]. In integrative group storytelling therapy, participants exhibited strong social integration, supported one another, displayed valuable life skills, and fostered positive self-worth and institutional loyalty during reminiscing and feedback sharing [21]. Consequently, participants felt supported and experienced a sense of belonging when performing digital group activities [22]. Community-based social participation relies on facilitators who organize activities and provide support in securing spaces and promoting events. Therefore, using a technology-based storytelling intervention to boost community-based social engagement can enhance community belonging.

The 3 main features of a digital storytelling application are multimedia material editing, memory recall, and story sharing. Some memory retrieval interventions may combine multiple functions of the interventions, such as people with MCI simultaneously reminiscing and sharing their recollections [23]. Challenges in using technology remain an unsolved issue; people with MCI typically exhibit lower digital literacy [24]. Most storytelling programs include training sessions before the intervention, such as storytelling classes [25], as mostly commercial, off-the-shelf technology is used to support digital storytelling [12], and these programs require digital literacy for multimedia editing and digital storytelling. Some studies required trained volunteers [26] or caregivers [23,27] to support people with dementia in developing stories. However, these settings are intended to enable people with MCI to adapt to existing environments and interactive tools rather than customizing the software to their specific needs (thus enhancing the usability of the storytelling application while simultaneously reducing learning costs). The applicability of these methods for memory retrieval among adults with MCI is currently not fully understood [28]. Therefore, our aim was to co-design a technology-based group memory retrieval intervention, which is a key form of digital storytelling intervention, to support social participation. We also investigated how to design user-friendly storytelling applications to manage digital resources and reduce learning costs for people with MCI.

Theoretical Framework
Co-design pertains to the collaborative involvement of people (users and stakeholders) in the design of a product or service [29]. Co-design workshops for interventions for people with MCI involve people with MCI, caregivers, and therapists in designing the intervention programs, providing unique perspectives. Involving users in design sessions helps engineers and researchers better understand requirements. A co-design approach would be helpful in identifying the requirements for
designing technology-based group memory retrieval interventions for people with MCI. In our study, the development of the storytelling prototype was guided by 2 behavioral and theoretical frameworks to raise participants’ awareness about their social health behaviors and tailor adequate actions for behavior change, namely the Behavior Change Wheel [30] and Theoretical Domains Framework (TDF). Beginning with the Behavior Change Wheel, a behavioral analysis enables intervention designers to select particular areas for exploration, thereby guiding the development of the digital storytelling approach. TDF deepens the understanding of psychological and behavioral factors influencing effective interventions. This integration aligned our application’s features with identified behavior change strategies, fostering a comprehensive approach to promoting social health. In addition, a communication strategy called MESSAGE (an acronym representing 7 key communication strategies, each encapsulated by its initial letter; within each strategy, specific suggestions are aimed at enhancing effective communication: maximizing attention [M], focusing on expression and body language [E], keeping it simple [S], providing support for their conversation [S], assisting with aids [A], getting their message [G], and encouraging and engaging in conversation [E]) [31] was used to support digital storytelling communication. Considering the critical importance of providing caregivers with effective tools to bolster memory and communication abilities in individuals with dementia [32], the implementation of specific communication strategies between people with MCI and the Huiyou application is paramount.

Methods

Ethical Considerations
The Swinburne University of Technology’s Human Research and Ethics Committee provided ethics approval for this research (20226525-11105) on September 26, 2022. All workshop participants provided written informed consent to participate and gave permission for their audio recordings and sketches to be used in publications. All methods were performed in accordance with relevant guidelines and regulations. We collected no identifying information from the research participants, such as their names or email addresses. Each participant was assigned a unique ID number to ensure their anonymity. In addition, we provided informed consent forms, including a project information sheet, to uphold their privacy rights.

Research Team
The research team is a collaborative ensemble of specialists in human-computer interaction, user experience, and design, each contributing their distinct expertise to the success of the project. AAM has an extensive background in design research and human-computer interaction, and WL has engaged in and performed research on user experience extensively. DZ is a doctor of philosophy student specializing in design.

Sample
A social work organization named Jingshilaonian, located in Beijing, China, expressed willingness to participate in this research. Situated within the Tiantongyuan community in Beijing, this institution specializes in providing psychosocial support, mental health services, and daily activity assistance to older adults, including those with MCI. Jingshilaonian played a vital role by assisting us with the recruitment and screening of people with MCI by administering the Montreal Cognitive Assessment, as developed by Nasreddine [33]. The inclusion criteria for people with MCI encompassed independent community dwelling, age >65 years, no visual or hearing impairments, and adequate reading ability. The exclusion criteria for people with MCI included significant neurological conditions, such as stroke or brain injury, because of potential confounding effects. Caregivers had no specific inclusion criteria. For therapists, the inclusion criteria were postgraduate qualifications and >3 years of experience in social interventions. To engage participants, we collaborated with the organization’s manager to distribute information sheets to people with MCI and their caregivers, inviting them to participate in the study. We co-designed a storytelling intervention with people with MCI (n=12), their caregivers (n=4), and therapists (n=5) in Beijing, China. The brainstorming stage and rough prototyping stage for the intervention were attended by people with MCI and caregivers in a group setting, and the user testing stage was attended by people with MCI and therapists. People with MCI (n=12) were aged 65 to 77 (mean 69.91, SD 4.20) years, and their average Montreal Cognitive Assessment-Chinese version score was 23.58 (SD 1.38). A total of 4 caregivers participated in the workshops. Table 1 summarizes the demographics of people with MCI.
Table 1. Demographics of participants with mild cognitive impairment (n=12).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Female</td>
<td>9 (75)</td>
</tr>
<tr>
<td><strong>Age (y), mean (SD)</strong></td>
<td>69.92 (4.20)</td>
</tr>
<tr>
<td><strong>Age (y), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>65-69</td>
<td>5 (42)</td>
</tr>
<tr>
<td>70-74</td>
<td>6 (50)</td>
</tr>
<tr>
<td>75-79</td>
<td>1 (8)</td>
</tr>
<tr>
<td><strong>Educational background, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>2 (17)</td>
</tr>
<tr>
<td>High school</td>
<td>3 (25)</td>
</tr>
<tr>
<td>Junior high school</td>
<td>7 (58)</td>
</tr>
</tbody>
</table>

Procedure

**Overview**

We adopted the co-design stages proposed by Robinson et al [34], namely scoping (stage 1), participatory design workshops (stage 1), and prototype development (stage 2). After developing the prototype, we conducted a usability evaluation (stage 3; Figure 1). The specific tools and procedures used can be found in the protocol paper [14].

Figure 1. The process of data collection. MCI: mild cognitive impairment; UEQ: User Experience Questionnaire.

Stage 1: Intervention Idea Brainstorming Workshops

In this initial phase, it took 45 minutes to uncover the essential aspects of significant social participation for people with MCI and their caregivers. This stage involved a sequence of 4 workshops designed to identify the most compelling benefits, role models, and barriers related to social participation. By prioritizing these aspects, we sought to deepen our understanding of the significance of social participation for people with MCI and enhance the quality of data for the subsequent phase. During each workshop, people with MCI and caregivers voted on the most compelling benefits and role models, and people with MCI selected 2 abilities that they wished to improve. These insights were critical for informing the design of our digital storytelling intervention.

Stage 2: Prototyping Workshops

Building on the insights gained from stage 1, we focused on collaborative brainstorming sessions involving people with MCI and their caregivers in stage 2. The goal was to explore the design and interaction logic of the interface through 60-minute workshops. Digital storytelling interventions involve recalling
past actions, events, and feelings using physical prompts. In addition to building on the insights gained from the literature review of digital storytelling interventions, we identified the main features of digital storytelling applications as story creation, memory retrieval, and story sharing [12]. Furthermore, we found that the themes of stories significantly impacted the storytelling experience [35]. Therefore, the research team proposed 4 primary properties for potential features of the digital storytelling intervention: identification of preferred storytelling themes, support for story material generation, memory retrieval, and story sharing. Participants were encouraged to describe, illustrate, or display sample interfaces on their mobile devices. Figures 2-6 were generated during the prototyping workshops by participants and facilitator DZ. The key objectives were to define the program’s main goals; propose relevant features, interactions, workflow, and interface elements; and create rough prototypes to visualize the intervention’s interfaces and interactions. The outcomes of this stage provided the foundation for designing the prototype of our digital storytelling application.

**Figure 2.** A sketch showing browsing and reminiscing topics and details.

**Figure 3.** A sketch of (A) material addition entry point indicated by a “+” icon and (B) adding material description of story materials.
Stage 3: User Testing

We created the Huiyou prototype based on the insights and ideas gathered from stages 1 and 2. To evaluate the application’s usability, we identified 22 specific tasks aligned with its features and interaction logic. These tasks were meticulously selected to ensure that the intervention’s primary objectives, including enhancing social connection and community participation for people with MCI, could be effectively achieved. Participants, including people with MCI and therapists, engaged in usability testing sessions, where they completed these tasks and provided feedback on their experiences. The usability test session was 45 minutes long. Feedback gathered from these sessions included task success rates, concerns, and preferences. We also collected data on individual task completion rates and task completion with facilitator assistance. In our usability testing, we used a comprehensive approach to evaluate the features and subtasks of the Huiyou intervention. For instance, in testing the
story sharing feature, participants were tasked with actions such as entering the speech mode, locating speech prompts, assessing speech length on the page, and identifying the screen projection button. If participants independently interacted with the interfaces successfully, the facilitator recorded it as task completion without assistance. However, if participants encountered difficulties in completing a task even after thoroughly examining the entire interface, the facilitator intervened by offering predefined tips, such as suggesting checking the bottom of the screen. In such cases where participants were able to successfully interact with the interfaces with the facilitator’s assistance, the facilitator recorded it as completion with facilitator assistance. Following this, we initiated an iterative design process to refine and enhance the user interfaces (UIs) based on participant feedback, ensuring the application’s usability and effectiveness.

Data Collection
Qualitative data were gathered through interviews, sketches, and focus group discussions at different stages of the research via audio recordings. In stage 1, focus group discussions were conducted to analyze the prioritized aspects of social participation. Stage 2 involved brainstorming activities during which participants generated sketches and discussed the potential features of Huiyou. In stage 3, participants’ feedback was collected through self-reporting during usability testing and structured interviews. It is important to note that the Chinese version of the User Experience Questionnaire (UEQ) had previously undergone psychological validation [36]. A total of 22 usability tasks were used to evaluate task completion, providing a comprehensive approach to gaining valuable insights into the user experience of Huiyou and effectively assessing its usability.

Data Analysis
Qualitative data analysis followed a thematic analysis approach [37,38]. Two individual researchers reviewed and coded the interview transcripts and organized the codes into themes using NVivo (version 12.0; Lumivero) [39]. Labels, such as “Stage 1-group 2-people with MCI number 2,” were used to distinguish different groups and participants.

In the design phase, the research team used the JiShiSheJi design concept to create the interface. After usability testing, a redesign phase was initiated, involving a detailed analysis of participant feedback and interview responses. These data were systematically summarized to identify improvement areas. An iterative design process was then used to enhance the UIs by incorporating participants’ feedback. This approach ensured that the final version of Huiyou aligned better with the needs and preferences of people with MCI.

Quantitative data analysis involved calculating completion rates using Excel (Microsoft Corp) to determine the average completion rate and feature rankings. The UEQ results were assessed using the UEQ Handbook [40], whereas the UEQ Data Analysis Tool Version 12 was used to compute mean values, SD, and CI for various usability dimensions of attractiveness, ease of use, efficiency, dependability, stimulation, and novelty. We also used Excel to calculate the mean values of task completion rates with and without assistance.

Results
The key findings from each stage of this study are summarized as follows.

Stage 1 Findings: Intervention Idea Brainstorming Workshops
Overview
Our exploration unfolded into 3 distinctive themes, each shedding light on a crucial aspect of participants’ experiences. We examined how social participation can be a powerful catalyst for enhancing physical health and fostering lifelong learning. Through the inspiring stories of confident older adults who wholeheartedly embraced challenges, our second theme showcased exemplars of unreserved social participation. Finally, we delved into the primary barriers people with MCI aimed to overcome, with a particular focus on memory, language expression, and the mastery of essential smartphone skills.

Enhancing Physical Health and Lifelong Learning Through Social Participation
Most participants with MCI (8/12, 67%) regarded social participation as an approach to maintaining physical health. They found that attending social activities can enhance their health status because they are less prone to catching colds when actively participating in social activities. People with MCI pay more attention to physical health; therefore, the benefits of physical health may attract people with MCI to participate in social activities.

Most participants with MCI (8/12, 67%) reported that engaging in activities with older adults positively impacted their physical well-being; these activities, such as exercise programs or outdoor group activities, helped improve their physical condition, increased their mobility, and supported their overall health. In addition, the social interaction involved in these activities afforded them a support system, along with accountability and the motivation to stay active and adopt a healthy lifestyle. In agreement with this, participants stated the following:

- Engaging with others and staying socially active can help improve my physical health. It encourages me to stay active, spend time with people, and motivates me to maintain a healthy lifestyle. [Stage 1, G2-P3]

Few participants with MCI (3/12, 25%) emphasized the value of lifelong learning and intellectual stimulation; participating in social activities provided these participants with a platform for communication, exploration, and the imparting of knowledge. By engaging in collaborative social activities and exchanging ideas with others, they were encouraged to continuously learn and discover new things. One of the participants commented the following:

- Engaging in social activities can improve my mental health by allowing me to interact with others, receive emotional support, reduce feelings of loneliness, and...
enhance positive emotional experiences. [Stage 1, G1-P2]

People With MCI Admire Peers Who Face Challenges Confidently

Most participants with MCI (9/12, 75%) admired older adults who displayed the confidence to attend various activities. Interacting with confident older adults allowed these participants to gain a new perspective on aging and life experiences; they were inspired to embrace challenges, overcome impediments, and approach life with a positive mindset:

They give me strength and inspiration, demonstrating passion and a positive attitude in caring for others, making me believe that I can still have a fulfilling and meaningful life even in challenging circumstances. [Stage 1, G3-P3]

People With MCI Want to Improve Their Memory and Language Deficit and Smartphone Skills

Table 2 indicates that most people with MCI wish to improve their memory and expression of language, as participants mentioned that memory affects them frequently and that language expression can affect communication with their grandchildren. Of the 12 participants with MCI, 7 (58%) prioritized the improvement of memory, as this affected them the most:

I have poor memory; as I spoke, I forgot where I was talking. Therefore, I want to improve it first. [Stage 1, G2-P2]

Table 2. The top 3 barriers identified (N=24).

<table>
<thead>
<tr>
<th>Barriers to overcome</th>
<th>Votes a, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memory ability</td>
<td>7 (29)</td>
</tr>
<tr>
<td>Language expression</td>
<td>5 (21)</td>
</tr>
<tr>
<td>Smartphone use</td>
<td>4 (17)</td>
</tr>
</tbody>
</table>

aEach person with mild cognitive impairment had 2 votes; in total, there were 24 votes from the 12 participants.

Stage 2 Findings: Prototyping Workshops

Overview

Following the analysis of the findings from stage 1, the research team believed that the creation of a group-based digital storytelling intervention would be beneficial for improving memory, language expression, and smartphone use skills, and eventually, it could enhance social connection and participation. We have outlined the findings regarding the 4 primary features of the digital storytelling application as follows.

Preferred Storytelling Themes for Sharing With Others

All the participants with MCI (12/12, 100%) stated that they would like to share recent activities, including hobbies (as well as the changes they perceived around them), social activities, and current politics. Hobbies were most commonly cited because the participants were eager to share them with others and had invested much time and effort in them. The participants with MCI reported the following:

I enjoy content related to technology, as well as driving. I like observing different cars; whenever a related topic arises, I cannot stop talking about it. [Stage 2, G2-P2]

Moreover, participants suggested receiving memory retrieval themes to ensure a well-defined topic to reminisce and allow them to gather relevant materials beforehand, expediting the process for smoother implementation.

Design of the Story Material Generation

After defining the storytelling themes, users could generate story materials by obtaining topics for memory retrieval, collecting materials, adding material descriptions, and integrating materials for further story sharing. The facilitator encouraged participants with MCI to create sketches; however,
only 4 (33%) of the 12 participants with MCI were inclined to do so (Figures 5 and 6). Consequently, the facilitator assisted those who were unwilling to draw in visualizing their ideas.

Figure 2 presents a sketch for browsing a topic for memory retrieval; on the left side of the page, there are a search bar for exploring events at the top and a list of events, including images, the name of the event, and its description at the bottom. The user navigates the list of events via a scrolling mechanism. Adjacent to each event entry is a registration button. The page on the right offers 2 options for displaying the activity. Option A displays only a text-based summary of the subject matter. By contrast, option B displayed a more significant activity schematic with introductory text provided underneath:

For example, for my son’s wedding, just get a few photos, just two or three. [Stage 2, G2-P3]

The material collection feature enhances content by providing explanatory details and enriches the collection by sourcing pictures from various internet platforms. Recording snippets of information further aids by capturing valuable insights. In addition to collecting the target content, participants also suggested recording fragmented information informally. Figure 3 presents a sketch of an entry point for the supplementary material at the bottom of the page, indicated by a plus icon, which reveals 3 input methods when clicked. The input methods include voice input, text input, and video input:

The function of voice recognition is very convenient. With just a press of a button, recording starts, and in the end, it can be converted into text for preservation, making it easy to view later. [Stage 2, G1-P2]

The adding material description feature aims to optimize the collection of data by associating individuals with each story (Figure 3); for instance, the developmental stages of the country, timeline, and the type of activities connecting people with MCI with the country’s developmental stages and arranging them chronologically and by different activities for easy navigation. Participants recommended the provision of a framework of memories to aid them in refining the story; for example, participants could select the time stage, such as the founding of the People’s Republic of China in 1949, and the stage of their life at the selected time, such as youth, middle-age, working, or retired. The choice of location is provided using a map. Finally, participants can apply labels for characters, including their avatars and names. One of the participants explained this as follows:

We can label the key characters that appear inside. [Stage 2, G3-P2]

I like to categorize these [materials] by age and objectively exist since I am willing to use classification methods that are specific, not easily confused, preferably objectively present, and not prone to misunderstanding. [Stage 2, G3-P3]

The integrating material feature aims to create a cohesive memory retrieval experience by uploading selected materials to the topic in question and adjusting sequences to promote a seamless and engaging narrative flow (Figure 4). The participants provided 2 options: option 1 displayed the candidate photos in the first row, and below, each theme was listed, with the corresponding material dragged and dropped into the theme; and option 2 allowed users to select the order of presentation after choosing the related theme. One of the participants with MCI commented the following:

After uploading these photos, we need to associate them with the story to be told based on different story themes. Before sharing, we could easily change the sequence of the materials. [Stage 2, G2-P1]

Memory Retrieval Activity Design Through Facilitation Strategies

The digital storytelling application uses various strategies to enhance the experience in the memory retrieval function. It has 3 features: a story presentation, group discussion, and story summary.

During story presentation, the digital storytelling application has cheat sheets for recall, minimizing interruptions by using timers, promoting interactive discussions, and creating permanent records. Mobile technology facilitates the easy capture of memories, forming meaningful connections with cherished life events. Preparing a cheat sheet with helpful tips can provide valuable guidance to ensure a smooth and engaging memory retrieval experience (as shown in Figure 5). The cheat sheet might present either keyword prompts or corresponding content. A cue word switches to the content of the corresponding material, depending on the selected content format:

Sometimes [my] memory is poor, and this software can provide a cheat sheet to prevent me from forgetting what I need to share at the moment. [Stage 2, G1-P2]

During the session, avoiding interruptions allowed participants to immerse themselves fully in their nostalgic journey, and a visual timer helped them manage their time effectively:

You can only limit the time. One is that most of the time, how many people cannot all stand on your own time for a few minutes. [Stage 2, G4-P2]

During a group discussion, the approach integrates memory retrieval with open discussion groups. Encouraging participants to pose questions by raising their hands fosters an interactive and dynamic environment (Figure 6). During the presentation, each participant has an avatar, and the Wi-Fi signal strength is displayed next to the avatar. If the participant has a question they wish to ask, they may raise their hand, and a small, raised hand icon is displayed next to their profile. These discussions can be informal, fostering a sense of engagement in shared experiences and collective memories:

Ask each other questions. This form is quite good. Ask each other because it is to discuss different opinions on a subject, and this is the best way. Yes, if there are no different opinions, this question may be a little biased. [Stage 2, G3-P3]

I want to ask questions. If you want to ask questions that I am interested in, I can also raise my hand and answer them. [Stage 2, G1-C1]
The story summary feature aims to make memory retrieval more tangible; facilitators can generate records or memories from the discussions, providing a lasting and meaningful resource for participants to cherish and revisit:

Currently, mobile phones are very convenient. When I take a group photo and see everyone’s photos, I can recall what happened at that time. I also like to keep a diary and can simply remember one or two sentences. [Stage 2, G4-C1]

**Story Sharing to a Broader Audience**

To share the records of their recollections with the desired person or group, participants suggested using WeChat (Tencent Holdings Limited), which allows for the convenient and efficient sharing of memories, ensuring that the memory retrieval experience can be cherished and enjoyed by those involved. Most participants with MCI (8/12, 67%) wished to share this experience with certain persons or groups, primarily through WeChat:

Based on different themes, [we can] send today’s activities to corresponding people, such as those with common interests or children. Share with them through WeChat and let them know the latest situation as well. [Stage 2, G2-P3]

However, one of the people with MCI showed no interest in sharing the content with others:

At this age, I’m not willing to share with more people; at least subjectively, I have no intention to please anyone. [Stage 2, G1-P2]

In summary, Table 3 outlines the key functionalities and features of the memory log system, along with valuable suggestions for their implementation. To generate materials, the inclusion of memory topics aids in focused preparation. Collecting materials might include adding explanations to materials, collecting pictures from the internet, and the informal recording of fragments of discussions. While adding material description, users could add details such as time, location, and key characters to the material. Users could integrate materials into each presentation by changing the sequence of materials. During memory retrieval, facilitators can benefit from cheat sheets, minimize interruptions, and use visual timers in presentations. Open discussions, hand raising, and interactive dialogues further enrich the experience. Summarizing these records enhanced memory retention. Sharing options, such as WeChat, extend the impact beyond the scope of the audience.

Table 3. The main features of Huiyou identified from stage 2 (ie, prototyping workshops).

<table>
<thead>
<tr>
<th>Functionalities and features</th>
<th>Suggestions to implement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generating story materials</strong></td>
<td></td>
</tr>
<tr>
<td>Obtain a topic for memory retrieval</td>
<td>• Inform the topic and collect relevant materials in advance</td>
</tr>
<tr>
<td>Collect materials</td>
<td>• Add a description to each record</td>
</tr>
<tr>
<td></td>
<td>• Collect pictures from others via the internet</td>
</tr>
<tr>
<td></td>
<td>• Casually record fragmented information</td>
</tr>
<tr>
<td>Add material description</td>
<td>• Add essential character to each material</td>
</tr>
<tr>
<td></td>
<td>• Link to the sequential steps of the country’s development</td>
</tr>
<tr>
<td></td>
<td>• Follow the timeline to show the content</td>
</tr>
<tr>
<td></td>
<td>• Categorize by different activities</td>
</tr>
<tr>
<td>Integrate materials</td>
<td>• Upload selected materials to the target topic</td>
</tr>
<tr>
<td></td>
<td>• Change the sequence of the materials</td>
</tr>
<tr>
<td><strong>Memory retrieval activity</strong></td>
<td></td>
</tr>
<tr>
<td>Story presentation</td>
<td>• Prepare a cheat sheet to obtain tips</td>
</tr>
<tr>
<td></td>
<td>• Avoid interruptions during the presentation</td>
</tr>
<tr>
<td></td>
<td>• Set visual timer</td>
</tr>
<tr>
<td>Group discussion</td>
<td>• Combine memories and discuss</td>
</tr>
<tr>
<td></td>
<td>• Raise hands to ask questions</td>
</tr>
<tr>
<td></td>
<td>• Open discussion</td>
</tr>
<tr>
<td>Story summary</td>
<td>• Generate a record or memory</td>
</tr>
<tr>
<td><strong>Sharing story</strong></td>
<td>• Share the record via WeChat (Tencent Holdings Limited)</td>
</tr>
</tbody>
</table>

**Developing the First Prototype of Huiyou**

The first digital prototype of Huiyou (Figure 7) was produced in JiShiSheJi [41], a web-based free software prototype design and development tool, following an iterative process of co-design workshops and feedback collected from the meetings. The software name 会友, Huiyou, derived from the Pinyin pronunciation, means “meeting new friends” in Chinese. It is
inspired by a classic quote from Confucius in the Analects: “A gentleman seeks friendship through literature and reinforces goodness through friendship.” This statement emphasizes the idea that individuals cultivate friendships through literary exchange and support virtue through companionship. The name reflects a positive vision of fostering social connections through literature, friendship, and benevolence. Users can leverage technological means through the software to expand their social circles, facilitating deeper communication and connections with others. The name embodies the social nature of the software and its goal of promoting friendship. Huiyou effectively stores memories from daily life and encourages people with MCI to reminisce and discuss favorable memories with new friends. Huiyou has 2 main features: supporting people with MCI to conduct self-reflection daily (preparing materials with cues) on certain topics and facilitating group memory retrieval (presenting a story and promoting a discussion with group members). Two innovative aspects of Huiyou are embedded memory retrieval for capturing daily life and its ability to collect recent, valuable memories. Another such feature is combined self-reflection and group reflection to enhance social interaction during the discussion, in addition to sharing and participating in social activities outside the home.

**Figure 7.** Huiyou prototype: home page.

As **Figure 8** illustrates, the process of registering an activity on the Huiyou application begins with the user navigating from the home page to the activity page, where they can browse and select the activity they wish to register for; they then upload the material and edit it on the active page. When an event is about to commence, the user enters the event details page from the home screen. Next, they enter the activity and encounter a start screen for analyzing memories. The activity may be recorded and shared with others. If the user has any issues with the software, they can click the help button to discover operation guidelines.
Stage 3 Findings: User Testing

Results From Task Analysis

Through this comprehensive usability testing stage, we identified minor areas of improvements in the interaction logic and interface layout to optimize user experience. In the 22 tasks, the participants scored an average success rate of 59% without assistance and 87% with assistance from the facilitator. Most challenges were identified to be associated with the direction of interaction; after knowing the direction, the success rate increased. Table 4 lists 5 tasks within the Huiyou digital storytelling application that demonstrated notably high usability, with completion rates exceeding 90% when performed individually. These tasks included “upload pictures” and “recording new materials,” both of which achieved perfect task completion rates of 100% with facilitator assistance. In addition, tasks such as “entering speech mode,” “find speech prompts,” and “insert a group photo after the event” also exhibited exceptional usability, with 92% individual task completion rates and 100% task completion rates with facilitator assistance.

This effectively communicates that within the Huiyou digital storytelling application, 5 tasks have a completion rate ≤75%, indicating the need for improvement. Tasks such as “add material description,” “browse registration activities,” and “switch the view of the activities” exhibited relatively lower individual task completion rates of 8%, whereas the presence of a facilitator notably enhanced completion rates to 58%. Similarly, the “judge whether the speech is too long on the page” and “enter the activity interface” tasks had initial completion rates of 25%, which significantly improved to 75% with facilitator assistance. These findings emphasize areas where the application’s UI and task guidance may benefit from refinement to ensure a more user-friendly and accessible experience, particularly for people with MCI.
Table 4. Task descriptions and completion rates of people with mild cognitive impairment (n=12).

<table>
<thead>
<tr>
<th>Features and task</th>
<th>Task completion without assistance, n (%)</th>
<th>Task completion with facilitator assistance, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Generating story materials</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upload pictures</td>
<td>11 (92)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>New materials for text input</td>
<td>6 (50)</td>
<td>10 (83)</td>
</tr>
<tr>
<td>New materials for voice input</td>
<td>7 (58)</td>
<td>10 (83)</td>
</tr>
<tr>
<td>Record new materials</td>
<td>12 (100)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Add material description</td>
<td>1 (8)</td>
<td>7 (58)</td>
</tr>
<tr>
<td>Add material tags</td>
<td>9 (75)</td>
<td>11 (92)</td>
</tr>
<tr>
<td>Edit material permissions</td>
<td>5 (42)</td>
<td>11 (92)</td>
</tr>
<tr>
<td>Enter my material</td>
<td>9 (75)</td>
<td>11 (92)</td>
</tr>
<tr>
<td>Change the sequence of materials</td>
<td>7 (58)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Browse registration activities</td>
<td>3 (25)</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Enter the activity interface</td>
<td>6 (50)</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Switch the view of the activities</td>
<td>1 (8)</td>
<td>7 (58)</td>
</tr>
<tr>
<td>Sign up for activities and add groups</td>
<td>5 (42)</td>
<td>9 (75)</td>
</tr>
<tr>
<td>View task progress</td>
<td>6 (50)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Enter task</td>
<td>10 (83)</td>
<td>11 (92)</td>
</tr>
<tr>
<td><strong>Memory retrieval activity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enter speech mode</td>
<td>11 (92)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Find speech prompts</td>
<td>12 (100)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Judge whether the speech is too long on the page</td>
<td>3 (25)</td>
<td>9 (75)</td>
</tr>
<tr>
<td>Find the screen projection button</td>
<td>9 (75)</td>
<td>12 (100)</td>
</tr>
<tr>
<td><strong>Sharing story</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insert a group photo after the event</td>
<td>11 (92)</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Use help features</td>
<td>4 (33)</td>
<td>10 (83)</td>
</tr>
</tbody>
</table>

**Results From the UEQ**

When comparing the ratings to the UEQ benchmark [36], it was noted that the application falls within the average and above range, suggesting that there is room for improvement. The combined ratings for all users, including both people with MCI and therapists organized by the value of each UEQ item can be seen in Figure 9, wherein the average score pertaining to attractiveness is 2.073 (SD 0.82), ease of use is 1.609 (SD 1.15), efficiency is 1.5 (SD 0.85), dependability is 1.875 (SD 0.76), stimulation is 1.813 (SD 0.96), and novelty is 1.703 (SD 1.33).

In the results obtained from the UEQ (7-point positive and negative scale; Figure 9), participants with MCI expressed favorable opinions about Huiyou, perceiving it as an enjoyable (mean 2.7, SD 0.6), supportive (mean 2.3, SD 0.7), clear (mean 2.5, SD 0.7), and friendly (mean 2.3, SD 0.8) application. However, they found learning challenging (mean 1.5, SD 1.9) and somewhat complex (mean 0.9, SD 2.1).
Results From Usability Testing

On the basis of the interviews conducted and feedback collected during usability testing, we identified 15 usability issues to improve (Table 5). These issues were linked to the features of the application, including the home page and navigation, sign-up activities, adding material descriptions, changing the sequence of selected materials, sharing stories, uploading material, UI and design, and picture gallery. The main change was made to the home page (resolving the first usability issue in Table 5), addressing the user feedback that there was a lot of information on the home page and that the activity entrance was not clear, making it difficult to locate it quickly. One of the participants with MCI mentioned the following:

*I’m not sure how to find the activities I want to participate in. The location of the activities should be more prominent. Those generated memories don’t necessarily have to be on the first page.* [Stage 3, P3]

Therefore, notable enhancements included moving activity records to a separate tab, thereby streamlining the navigation and organization processes. In addition, more activity entries were added to the home page, providing users with a broader range of options. We followed all recommendations to redesign the Huiyou application.
Table 5. Usability issues and recommendations.

<table>
<thead>
<tr>
<th>Features and issues</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Home page and navigation</strong></td>
<td></td>
</tr>
</tbody>
</table>
| There is a lot of information on the home page, and the activity entrance is not clear | • Show the activities list. Each activity may have a different status: ongoing activity (in a dark green rounded rectangle) and waiting to start (in a gray rounded rectangle)  
• Move the entrance of the “casually record fragmented information” feature from being suspended in the lower right corner to being fixed in the lower center |
| Category activity is hard to find | • Move activity records to a separate tab  
• Add more entries on the home page |
| **Sign-up activities** | |
| Activity registration setup issues | • Display activity registration first, followed by recent activities |
| **Adding material description** | |
| Unable to understand the edit permissions | • Change the heading to “invite others to edit” |
| Unable to find materials and view photos that have already been uploaded | • Integrate the image library and add tags  
• Allow easy switching and searching for images |
| Unable to understand content edit functions | • Change to content description, integrating the description on 1 page, including time, location, characters, events, and others  
• User should be able to add pictures’ tags from material, including family, friend, sightseeing, group photo, and item |
| **Changing the sequence of selected materials** | |
| Unable to sort materials | • Provide operating instructions and make the font color of the instructions more eye-catching  
• Provide support for 2 sorting methods: drag and click |
| **Sharing story** | |
| The “share the screen” button is too small | • Make the screen projection button larger and highlight the color more prominently |
| Face-to-face discussions require no “raise hand” button | • Remove the “raise hand” button for inquiries in face-to-face discussions |
| **Uploading material** | |
| No need to find internet resources | • Delete the option of “upload material through internet” |
| **User interface and design and picture gallery** | |
| Need colorful and simple icons | • Update icons to be colorful, eye-catching, and more solid |
| The return icon is not prominent | • Increase the size of the return icon  
• Change the dark background of the “help” icon to a lighter background |
| Keyboard input and voice input | • Default voice input |
| The picture gallery contains too much information | • Reduce the number of images displayed in galleries and preset some images in advance |

**Discussion**

**Principal Findings**

This study outlines the process and outcomes of co-designing and prototyping Huiyou, a digital storytelling application intended to facilitate social engagement and enhance the cognitive well-being of people with MCI in community-based settings. Insights collected from usability testing shed light on both strengths and areas for improvement in terms of user-friendliness and accessibility within the application’s design and functionality. Furthermore, some functions have been simplified, including the path for uploading material without internet resources. People with MCI may take 15 minutes to collect the materials they want to present and 10 minutes to present their stories. Each story is original, proposed by people with MCI, and they may ask for the support of volunteers or caregivers. The storytelling process has 2 phases: preparation of materials and memory retrieval. Emphasizing the recollection
of recent memories was found to encourage social engagement and foster a sense of belonging. Huiyou, the storytelling application used in this study, facilitated material generation by providing preset content prompts and allowing independent material collection, setting it apart from other interventions. During the memory retrieval activity, participants used personalized cues and multimedia elements, triggering meaningful conversations and connections. This study’s findings suggested that storytelling themes for people with MCI should revolve around recent positive experiences and significant life periods. Notably, the application scored highly in attractiveness, dependability, stimulation, and novelty, although it required ease of use and efficiency enhancements. Recommendations for interface design included emphasizing crucial elements, minimizing cognitive complexity, and streamlining information presentation for improved user accessibility and experience.

Comparison With Prior Work
This research highlights the importance of selecting storytelling themes that evoke recent positive experiences for people with MCI in China. The emphasis on significant life stages and changes, excluding marital experiences, is supported by existing literature [37]. Surprisingly, the study observed a keen interest among older male adults in political subjects, deviating from expectations based on previous research on political engagement among older Chinese adults [42]. It is proposed that prioritizing recent positive memories through Huiyou can encourage social engagement, fostering a sense of belonging and participation among people with MCI. Huiyou facilitates the storytelling process for people with MCI through self-collected materials rather than preset content prompts [43,44], engaging them to share their stories. According to the Capability, Opportunity, and Motivation–Behavior (COM-B) system, Huiyou aims to enhance psychological capability, create social opportunities, organize social activities regularly, and reinforce reflective and automatic motivations. In addition, the following TDF components were embedded in the prototype: skills, social role and identity, beliefs about capabilities, goals, memory, attention and decision processes, and social influences. For example, Huiyou incorporates behavior change strategies in the TDF that involve social influence [45], such as group discussions that monitor group progress. As the therapists said, these design features can foster peer pressure, enhancing adherence to task completion and effectively facilitating the establishment of social connections. During group discussions, Huiyou implements the MESSAGE communication strategy to involve people with MCI actively. For instance, it allows users to add notes to materials, which are then displayed on the screen during memory retrieval. Unlike other interventions, Huiyou enables users to collect and arrange materials themselves, enhancing their sense of accomplishment and reducing the need for external support. The application’s approach of facilitating material generation and arranging sequences aims to stimulate positive memories and encourage active engagement during the memory retrieval process. Therefore, Huiyou supports people with MCI in arranging the display sequences of materials, as OurStory does [23]. In memory retrieval activities, a substantial number of stimuli are prepared in advance and presented randomly to prompt older adults to narrate stories [46]. However, this approach relies on the divergent thinking abilities of older adults. Unfortunately, this method does not support the possibility of multiple stimuli coming together to form a more complete story. Unlike older adults with dementia, people with MCI possess the autonomy to select cues for their storytelling, fostering interpersonal connections and evoking positive emotions [47]. Personal topics serve as effective memory tests, enhancing storytelling and social memory [48]. Huiyou uses visual cues, music, and various technological platforms to stimulate memory retrieval and trigger discussions among older adults. By encouraging the recall of recent positive memories and promoting group discussions [49], Huiyou enhances the confidence of people with MCI, empowering them to actively participate in social interactions and community activities. Group discussions not only provide opportunities for self-expression but also foster reflection and inspiration from shared social experiences. Sharing recent memories via Huiyou fosters dialogue and active social engagement among individuals with MCI. They can either share their stories during group reminiscing sessions or record and distribute their memories to family and friends. Unlike some interventions that lack a structured approach to memory retrieval [50], Huiyou allows for the systematic recording and sharing of stories [51], primarily using WeChat as the chosen platform for its broad reach among friends and family.

Usability Issues of Huiyou
On the basis of the UEQ benchmark results diagram, Huiyou earned a place in the “excellent” category [36]. These insights serve as invaluable pointers for refining the user experience of the application. People with MCI typically adhere to a top-to-bottom, left-to-right reading pattern [52], emphasizing the importance of placing essential interactive elements at the center or the top of the screen. Furthermore, adopting a more noticeable design approach, such as incorporating colored buttons, may be beneficial. The cognitive complexity experienced by people is influenced by the quantity of content displayed on a single screen. Even subtle variations in interface design demand additional information processing time for people with MCI. To mitigate cognitive strain, it is advisable to minimize the amount of information displayed on each screen and segment tasks into more manageable steps.

Limitations
One limitation was the small number of participants with MCI throughout the study. To mitigate the impact of a small sample size, the research team invited the same participants to engage in multiple stages of the study, gathering diverse research data, including interview outcomes, hand-drawn interfaces, and usability task data, and introduced the perspective of caregivers of people with MCI. Reusing the participants was an efficient approach to prototype development and, to some extent, mitigated the impact on people with MCI as well as clinician time. Another limitation was that the Huiyou prototype was redesigned based on the participants’ feedback; however, in this study, we did not evaluate the refined prototype. In the future, field testing with people with MCI will be used to measure the effectiveness of the tool in improving social participation.
Conclusions
We described the co-design processes of developing a digital storytelling intervention, Huiyou, in collaboration with people with MCI and caregivers. We then evaluated the user experience of the application based on the feedback of people with MCI, caregivers, and therapists. Huiyou incorporates story-sharing themes that align with the needs of people with MCI in China. These themes foster common topics and evoke positive emotions without delving into excessive privacy. Unlike traditional reminiscence therapy, it is confined to memory enhancement during intervention sessions, neglecting other potential intervention times. Huiyou transforms reminiscence into an everyday activity that individuals can engage in at their own convenience. This provides more opportunities to collect and cherish fond memories. Huiyou excelled in usability testing, earning an “excellent” rating in the UEQ benchmark for attractiveness, reliability, stimulation, and novelty. However, there is room for improvement in accessibility and efficiency. By combining social participation with the fostering of relationships and the stimulus to contact friends, the application not only promotes individual well-being but also meaningful social interactions and maintains vital relationships for people with MCI.

Acknowledgments
The authors would like to thank all the staff, people with MCI, and caregivers who were involved in this project.

Authors' Contributions
AAM and DZ conceptualized this study. DZ collected and analyzed the data. AAM directed several rounds of feedback and contributed significantly to the manuscript’s writing and revision. AAM and WL supervised the project. All the authors examined and approved the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Higher resolution version of Figure 8. [PNG File, 1075 KB - aging_v71e54138_app1.png]

References


34. Robinson L, Brittain K, Lindsay S, Jackson D, Olivier P. Keeping In Touch Everyday (KITTE) project: developing assistive technologies with people with dementia and their carers to promote independence. Int Psychogeriatr 2009 Feb;21(03):494. [doi: 10.1017/s1041610209008448]


41. Home page. Gees Al. URL: http://gees.ai/ [accessed 2023-12-25]

42. Xie B, Jaeger PT. Older adults and political participation on the internet: a cross-cultural comparison of the USA and China. JMIR Aging 2024 | vol. 7 | e54138 | p.173https://aging.jmir.org/2024/1/e54138 (page number not for citation purposes)


47. Robinson L, Brittain K, Lindsay S, Jackson D, Olivier P. Keeping In Touch Everyday (KITTE) project: developing assistive technologies with people with dementia and their carers to promote independence. Int Psychogeriatr 2009 Feb;21(03):494. [doi: 10.1017/s1041610209008448]


Abbreviations

COM-B: Capability, Opportunity, and Motivation–Behavior
MCI: mild cognitive impairment
MESSAGE: maximizing attention, focusing on expression and body language, keeping it simple, providing support for their conversation, assisting with aids, getting their message, and encouraging and engaging in conversation
TDF: Theoretical Domains Framework

https://aging.jmir.org/2024/1/e54138 JMIR Aging 2024 | vol. 7 | e54138 | p.173
(page number not for citation purposes)
Experiences of Older Adults, Physiotherapists, and Aged Care Staff in the TOP UP Telephysiotherapy Program: Interview Study of the TOP UP Interventions

Rik Dawson1; Heidi Gilchrist1; Marina Pinheiro1; Karn Nelson2; Nina Bowes3; Cathie Sherrington1; Abby Haynes1

1Institute for Musculoskeletal Health, Sydney Musculoskeletal Health, Sydney Local Health District, The University of Sydney, Camperdown, Australia
2Whiddon, Sydney, Australia
3Uniting AgeWell, Melbourne, Australia

Corresponding Author:
Rik Dawson
Institute for Musculoskeletal Health, Sydney Musculoskeletal Health
Sydney Local Health District
The University of Sydney
PO Box M179, Missenden Road
Missenden Road
Camperdown, 2050
Australia
Phone: 61 403895186
Email: rik.dawson@sydney.edu.au

Abstract

Background: Telehealth provides opportunities for older adults to access health care. However, limited research exists on the use of telehealth within aged care services, particularly regarding physiotherapy-led fall prevention and mobility programs. Understanding the experiences and interactions of older adults, physiotherapists, and aged care service providers is crucial for the scale-up and sustainability of such essential programs. The TOP UP study, a hybrid type 1 effectiveness-implementation randomized controlled trial in aged care, used a supported multidisciplinary telephysiotherapy model to motivate older adults to engage in exercises to improve mobility and reduce falls.

Objective: This qualitative substudy aims to achieve 2 primary objectives: to describe the experiences and acceptability of the TOP UP intervention for older people, physiotherapists, and aged care support workers and managers and to gain an in-depth understanding of program implementation.

Methods: A purposive recruitment strategy was used to select 18 older adults who participated in the TOP UP intervention, ensuring variation in age, gender, residential status (home or residential aged care), geographic location, and cognitive levels. In addition, 7 physiotherapists, 8 aged care support workers, and 6 managers from 7 different aged care provider partners participated in this study. Semistructured interviews were conducted to explore stakeholders’ experiences with the TOP UP program, gather suggestions for improvement, and obtain insights for the future implementation of similar telephysiotherapy programs. The interview framework and coding processes were informed by behavior changes and implementation frameworks. Data were analyzed using an abductive approach, informed by 2 behavioral change theories (Capability, Opportunity, Motivation, and Behavior Model and Self-Determination Theory) and the Nonadoption, Abandonment and Challenges to the Scale-Up, Spread and Sustainability of Health and Care Technologies framework.

Results: All participants (n=39) reported high levels of acceptability for the TOP UP program and cited multiple perceived benefits. The thematic analysis generated 6 main themes: telephysiotherapy expands opportunity; tailored physiotherapy care with local support enhances motivation; engaging, older adult–friendly educational resources build capability; flexible reablement approach fosters autonomy; telephysiotherapy is safe, effective, and acceptable for many; and organizational commitment is required to embed telehealth. The motivation to exercise was enhanced by Zoom’s convenience, use of tailored web-based exercise resources, and companionable local support.

Conclusions: This study highlights the inherent value of telephysiotherapy in aged care, emphasizing the need for investment in staff training, local support, and older adult–friendly resources in future telephysiotherapy iterations. TOP UP represents a...
convenient and flexible web-based care model that empowers many older adults to receive sustainable, high-quality care precisely when and where they need it.

**Trial Registration:** Australian New Zealand Clinical Trials Registry (ANZCTR) ACTRN 1261000734864; https://anzctr.org.au/Trial/Registration/TrialReview.aspx?ACTRN=12621000734864

*(JMIR Aging 2024;7:e53010) doi:10.2196/53010*

**KEYWORDS**

physiotherapy; telehealth; telephysiotherapy; exercise; aged care; qualitative methods; behavior change; technology; virtual care

### Introduction

#### Background

The proportion of older people in the population is increasing worldwide. From 2020 to 2050, the number of older people aged ≥60 years will double to 2.1 billion, representing 22% of the world’s population [1]. By 2050, the number of people aged ≥80 years is expected to triple to 426 million [1]. Older people experience poor mobility and higher rates of falls than younger people, leading to significant morbidity, mortality, and poor quality of life [2–4]. Poor mobility and falls are 2 of the biggest cost drivers in hospital and aged care services [2,5]. In 2021, a total of 10.7% of people aged ≥65 years living in the Organization for Economic Co-operation and Development countries received long-term care, either at home or in care facilities, costing these governments 1.5% of the gross domestic product [6,7]. Aged care spending has increased over the last 15 years in most Organization for Economic Co-operation and Development countries, and population aging will continue to increase the demand on stretched health care systems [7].

Strong evidence supports the effectiveness and cost-effectiveness of physiotherapy-led exercise programs for enhancing mobility and reducing falls in aged care settings [8]. However, the Australian Royal Commission into Aged Care highlighted significant barriers to accessing mobility-promoting and fall prevention interventions delivered by allied health professionals, such as physiotherapists [9]. Qualitative evidence suggests that (referred and defined in this manuscript as telephysiotherapy) could be a feasible, acceptable, and effective approach for delivering mobility and fall prevention programs to older adults living in the community [10]. Telephysiotherapy could be particularly advantageous in increasing access and convenience for people with travel constraints and mobility limitations or who live in regional and remote areas [11].

Telerehabilitation that has provided telephysiotherapy has been found to have similar effectiveness compared with in-person rehabilitation services for community-dwelling older people, and it shows no increased risk of adverse events [12]. However, there is no evidence supporting the effectiveness, cost-effectiveness, and implementation feasibility of telephysiotherapy for improving mobility, reducing falls, and enhancing the quality of life for older adults receiving aged care services in their homes or residential aged care.

Telehealth is currently being used in aged care, but there is limited guidance on how best to implement it [13]. Hybrid effectiveness and implementation research has been shown to accelerate research translation into clinical practice [14]. Implementation research explores the experience of a complex intervention such as telehealth and its relationship to other factors, such as intervention engagement and adherence, perceived effectiveness, acceptability, and self-efficacy, which can support implementation translation [15].

#### The TOP UP Trial

TOP UP is designed to provide a scalable solution for delivering physiotherapy exercise interventions via telehealth to improve mobility, reduce falls, and enhance the quality of life in aged care. The TOP UP program was developed in collaboration with our aged care partners, physiotherapists, and aged care service users and their caregivers. A series of workshops identified potential facilitators to improve older people’s engagement with technology and motivation to exercise. The program’s co-design was also influenced by behavior change models, such as Self-Determination Theory (SDT) [16] and the capability, opportunity, motivation, and behavior (COM-B) framework [17].

TOP UP is investigating synchronous and asynchronous care to optimize both personalized health care and self-directed exercise [18] in aged care settings. It involves the delivery of real-time physiotherapy assessments through videoconferencing (synchronous telehealth) using the Zoom app (Zoom Video Communications Inc) by older adults receiving aged care services at home or in residential care. These service users are given access to evidence-based exercise videos on the TOP UP website and the StandingTall app (asynchronous telehealth) to support their exercise program. Each participant has the weekly support of a trained aged care worker to help them access the Zoom app and follow the exercise program. The outcomes being measured include effectiveness (mobility, falls, and quality of life), cost-effectiveness, and implementation measures (acceptability, reach, fidelity, dose delivered, and adoption).

The program is being tested in a hybrid type I effectiveness-implementation randomized controlled trial. Older adults were screened by their aged care service providers. Eligibility criteria included the age of older people (≥65 years); possessing sufficient physical, sensory, cognitive, and English language skills to participate; and having individual consent or consent from the person responsible. Those with terminal or unstable illness, with severe dementia, having participated in a similar physiotherapy program in the previous year, or being unable to walk 10 m were excluded from the study. A total of 242 participants were recruited from a screening pool of 1348 aged care service users (older people).

A total of 242 participants (120 per group) will provide 80% power to detect a 0.9 point between-group difference in 12-point
Short Physical Performance Battery test scores at 6 months (assuming SD 2.8, \( P=.05 \), and 20% dropouts) [19]. A 0.5-point between-group difference in the Short Physical Performance Battery test was considered clinically significant. This sample size was expected to be sufficient to detect between-group differences of 10% to 15% for the secondary outcome measures. Quantitative data analysis is expected to be completed in 2024.

Participants randomized to the intervention group received 10 videoconference physiotherapy sessions over 6 months using the Zoom app and received an individualized balance and strength exercise program. These exercise programs are based on the World Health Organization 2020 guidelines on physical activity and sedentary behavior [20] and the Otago exercise programs [21]. Existing aged care support staff, called “coaches,” have been trained to supervise participants to access the technology and provide “hands-on” exercise support once per week with the assistance of exercise videos designed by the research team. The waitlist control group receives a 3-month version of the program once the intervention period at each site is completed.

To inform the successful development of programs such as TOP UP, it is essential to examine not just if but how and why TOP UP worked (or not) and what strategies could best improve it. The aim of this paper was to use interview data to provide detailed insights into the experiences of older people, physiotherapists, coaches, and aged care managers with the telephysiotherapy intervention. The objective was to understand how contextual factors mediate the delivery of the TOP UP program and to produce transferable lessons for the potential use of future telephysiotherapy in aged care [22].

**Methods**

**Study Design and Context**

This study used a qualitative, descriptive approach through semistructured one-on-one interviews [23]. Qualitative description is increasingly used in conjunction with effectiveness and implementation trials and aims to present a straightforward description of participants’ experiences [24]. The analysis is grounded in the participants’ own words, making the results accessible to vulnerable groups, valid, highly translatable, and useful for refining interventions [25]. Qualitative description sits within a constructivist paradigm, considers multiple meanings, and recognizes that the research process is never neutral [26]. To strengthen the research rigor, we included triangulated data sources (by drawing on perspectives of different stakeholder groups) and a reflective discussion of emergent findings among the multidisciplinary research team [27].

**Conceptual Framework**

We used 2 behavior change theories, COM-B and SDT, and the Nonadoption, Abandonment and Challenges to the Scale-Up, Spread and Sustainability of Health and Care Technologies (NASSS) framework to provide a conceptual “lens” to inform data collection and analysis [28]. The COM-B model of behavior change proposes that to engage in a behavior such as exercise (B), a person must be physically and psychologically capable (C) and have the opportunity (O) to engage in the behavior, as well as the motivation to do so (M). COM-B simplifies complex factors and recognizes that to modify behavior, we need to address at least one of these components [17]. SDT focuses on the motivation underpinning behavior change, positing that effective programs must support autonomy, competency, and relatedness [16]. The NASSS framework is an evidence-based, theory-informed, and pragmatic framework that can help predict and evaluate the success of a technology-supported health program. It consolidates multiple implementation frameworks, targeting key issues relating to the implementation and uptake of telehealth at the microlevel of individual staff and consumers, the mesolevel challenges of organizational engagement and adoption, and macrolevel policy and regulatory factors (Figure 1) [28].
Recruitment and Data Collection

At the initial TOP UP recruitment, all aged care service users, coaches, physiotherapists, and aged care managers received an information sheet inviting them to participate in an interview for this qualitative study. After participants read the informational letter and confirmed their interest in participating in an interview, they received an informed consent letter to be signed by themselves or their person responsible before the interview appointment. A list of potential aged care service users and their coaches and physiotherapists was created in consultation with 3 of our aged care partners (Ashfield Baptist Homes, Whiddon, and Uniting AgeWell).

A purposive recruitment strategy was used to select 18 older adults who participated in the TOP UP intervention, ensuring variation in age, gender, residential status (home or residential aged care), geographic location, and cognitive levels. All 39 participants contacted agreed verbally and in writing before and on the day of the Zoom interview. None of the participants declined to participate in the interviews. The interviews were conducted 3 to 6 months after the interviewees commenced the program. Recruitment was stopped at the point when data adequacy had been reached, that is, when we judged that we had sufficient rich data across our purposive sample with which to answer our research questions [29].

The interview guide was created in consultation with the wider research team and representatives from our aged care partners (Multimedia Appendix 1). Interview questions targeted concepts from the COM-B model, SDT, and NASSS framework (described earlier). Specific questions explored the relative value of the different components of TOP UP (eg, Zoom, exercise videos, and the level of support provided). Questions regarding its implementation and effectiveness were also included. We also asked interviewees to identify local and potential system-wide barriers and facilitators to the successful adoption of telephysiotherapy, such as TOP UP and other programs aimed at delivering fall prevention and mobility programs in aged care. We used open-ended questions and active listening to confirm our understanding of the interviewees’ perspectives. RD and KN conducted the interviews individually on Zoom. They were
involved in the delivery of the program, so to reduce sociability bias, interviewees were encouraged to critique the TOP UP program and its implementation to identify improvements [30].

Aged care participants had a family member or someone familiar to them from their aged care organization that was not their coach to assist them with connection to the Zoom app and to support them through the interview. The participants were reminded that they could stop the interview at any time. No repeat interviews were conducted. RD and KN completed memos after the interviews and met to discuss the data and the emerging thematic content. Interview audio recordings were automatically transcribed using Zoom’s free transcription service, and transcripts were corrected by RD. Transcripts were not returned to the participants.

Data Analysis
The transcripts and interview field notes were uploaded to NVivo 12 (Lumivero) for data management and coding [31]. The transcripts were coded by one researcher (RD) using an abductive analytical approach. RD drafted an initial thematic framework drawing on emergent themes in the data and was informed by domains from the NASSS framework. In total, 2 researchers (RD and AH) tested and refined the coding framework on 2 manuscripts, adding codes and modifying existing codes from inductively identified concepts in the data. RD coded the remaining data. RD, CS, and AH met regularly to discuss emergent codes and themes.

Recurrent themes were generated from reading across the coded data and reviewed against concepts from SDT [16] and the COM-B framework [17,29] to understand how aspects of the TOP UP program influenced exercise engagement. An early overview of the findings was discussed with all coauthors and our consumer representative to explore a wider range of possible thematic interpretations and to help ensure that we had answered our research questions, including considering the implications of our findings. Disagreements were resolved through discussion. The criteria for reporting qualitative research was used as a reporting checklist (Multimedia Appendix 2) [32].

Ethical Considerations
Ethics approval for this qualitative substudy was included in the TOP UP study approval granted by the Ethics Review Committee at the Sydney Local Health District Research Ethics and Governance Office, Concord, Australia (approval number CH62/6/2021-009). The trial was registered with the Australian and New Zealand Clinical Trials Registry (ACTRN 1261000734864).

Results
Participants
In total, 39 people participated in semistructured interviews: 18 (46%) aged care service users who completed the TOP UP program, 7 (18%) aged care physiotherapists, 8 (21%) coaches, and 6 (15%) aged care managers. Interviews took an average of 19 (range 8-53) minutes. These service users were aged from 70 to 93 (median 87.5) years at the start of the intervention; 11 (61%) were female and 7 (39%) were male; 11 (61%) used a 4-wheeled walking frame to walk and 7 (39%) did not need a walking aid to walk; 7 (39%) lived in metropolitan cities in New South Wales and 11 (61%) lived in rural or remote areas in New South Wales and Victoria; 6 (33%) had mild to moderate cognitive impairment and 12 (77%) had no cognitive impairment; all had multiple comorbidities (median 7, range 2-11); and 10 (56%; median 1, range 1-7) had one or more falls in the last 12 months. The median Technology Readiness Index score was 2 out of 5 (range 1-3.9), which classified aged care users as technology avoiders, people who tend to have a high degree of resistance and a low degree of motivation to use technology [33]. A total of 4 (22%) aged care service users had used phone-based telehealth before with their general practitioner, but none had used a videoconferencing app such as Zoom before the study or used telehealth to receive physiotherapy.

Of the 7 physiotherapists interviewed, 4 were based in metropolitan areas and 3 were based in rural areas. A total of 4 coaches supported aged care service users from residential aged care, and 4 coaches supported aged care service users from home aged care. In addition, 1 aged care manager worked at a remote residential aged care site, 1 managed a rural residential aged care site, and 4 were home care managers from rural areas.

Main Findings
Overview
Our qualitative analysis revealed that all interviewees found the TOP UP program to be acceptable and would recommend similar telephysiotherapy programs to other older people receiving aged care services. Thematic analysis generated 6 key themes related to the experiences of TOP UP. We also compiled evidence of these experiences to identify and manage emergent possibilities, uncertainties, and interdependence that could guide the adoption of telephysiotherapy in aged care using the NASSS framework. Quotations were used to illustrate each theme. We annotated the quotes for anonymity with aged care service users referred to as P1, P2, and so on, and other stakeholders are descriptively described.

Theme 1: Telephysiotherapy Expands Opportunity
Theme 1 highlights the expanded opportunities for accessing the physiotherapy that TOP UP provided. TOP UP minimized barriers to physiotherapy access related to travel and associated costs:

Travel in country areas is just too hard and having telehealth in the home makes it so easy to do. I can’t do a 70 km round trip – it is too expensive. [P10]

This was also echoed by service users whose significant disabilities created access barriers:

Because of my health, there’s no way I can go out to see a physio. One, I’ve got to get someone to take me, like a relation, or pay someone to take you, it’s not practical. It’s hard to park anywhere near the physio, you’ve got to walk, so by the time you get to the physio you’re exhausted. [P1]

All the physiotherapists and aged care service managers indicated that telehealth could deliver physiotherapy care
efficiently, improving opportunities for older adults receiving aged care services to receive physiotherapy where and when they need it:

Some people need to be able to see a physio quickly, and we can provide telehealth services quickly, it is so efficient. [Home care physiotherapist, metro]

TOP UP had the greatest impact on rural and remote services, especially in areas where telehealth has the potential to address chronic health inequity issues related to workforce shortages:

Our town has a physio that visits once a month. Recently, one of our residents had a fall so I called the clinic and found out that we can’t get an appointment to take our resident to see a physio for 6 weeks, if someone has a fall like this, we just can’t wait six weeks. Telehealth really helps us. [Residential manager, remote]

Theme 2: Tailored Physiotherapy With Local Support Enhances Motivation

Regular local support was identified by all interviewees as important for enhancing older people’s confidence to try the exercise program and to support their motivation to “stick with” the program, including coping with TOP UP’s increasing challenge over time. Many interviewees across the 4 stakeholder groups explained that it was not just physical and technological support that the coaches offered (eg, providing stand-by assistance while performing balance exercises and managing the Zoom app) but also companionship and emotional support:

I can’t get out much and I began to look forward to the weekly session with my care worker as I really appreciated the support, she gave me to do something positive for my health. My coach understood what was going on in my life and she gave me the confidence to keep doing the exercises. [P12]

All the physiotherapists interviewed indicated that the coaches’ “hands-on” support was vital to the success of the program, as it helped to build capability and confidence. Importantly, TOP UP was a tailored program in which physiotherapists were able to modify exercises according to the individual needs of each service user, mirroring the person-centered approach typical in-person physiotherapy sessions. This tailored approach was particularly important to aged care service users whose health changes required program adaption:

My physio understood that I needed a break when I had surgery, but was ready for me when I got back home and quickly helped me regain the fitness that I had lost in the hospital. [P18]

The TOP UP program used technology and behavior change techniques to maximize program adherence. Zoom provided physiotherapists with a platform to deliver individualized real-time health coaching and goal setting, which has been shown in the literature to increase participant adherence [34]. Interview data indicated that these techniques were being used across program implementation, consistent with the behavior change theory:

Physiotherapists start with building external motivation by setting goals, by encouraging them, and highlighting their progress we help them develop internal motivation to keep going. If we can motivate them internally, half the job is done, and exercise will become a routine and a lifestyle habit. [Home care physiotherapist, metro]

Further evidence emerged that these techniques were having the desired effect:

Their motivation seemed to improve when they reached their goals, and they wanted to keep on trying. Their motivation is the most important thing. [Residential coach, metro]

However, it was the combination of live tailored physiotherapy with enthusiastic and companionable local support that seemed to both develop confidence and underpin motivation:

The individual sessions on Zoom were important, so I could ask some questions about how I was doing, and having my physio give me some individual feedback was important for my confidence to keep exercising. My coach has made it possible, and her support has been great, she is so lively, and she exercises with me which makes it fun, we had such a laugh, she keeps me motivated, and she takes the monotony out of it. If you are not having fun, it is not worth it. [P10]

The previous quote also highlights the vital role of enjoyment in exercise and how this can be enhanced with a trained support worker acting as a coach. This may be especially important for engagement in the TOP UP program, given that many of the physiotherapists interviewed suggested that telehealth requires more time to develop a therapeutic alliance. Therapeutic alliance refers to how people experience the empathy of clinicians, and research shows that a strong therapeutic alliance is connected to positive treatment adherence and results in physical therapy [35]:

It’s not until you get to the fourth or fifth telehealth session that people start getting to really know you and feel like you can be an advocate for them. I think that telehealth does allow for a personal connection which adds to exercise adherence. [Physiotherapist, home care, rural]

TOP UP physiotherapists and coaches’ person-centered approach to goal setting, highlighting progress during the program and celebrating any achievements, seemed to enhance the aged care service user’s motivation to exercise:

TOP UP helped me care about my future before I just didn’t care. I loved the way the physio explained things to me carefully, so I understood. I really appreciate having the support worker exercise with me and reinforce what to do, and how to do the exercises. All of this made me feel like I mattered and now I can walk further, do my shopping, which is a big improvement. [P12]
Theme 3: Engaging, Older Adult–Friendly Educational Resources Build Capability

Interviewees expressed enthusiasm for the instructional videos that were designed to support high-quality independent exercises throughout the program. The videos incorporated exercises modeled by an older person, slow-paced dialog in a warm conversational style, natural lighting to maximize visibility, minimal visual distractions, and gentle humor, all of which seemed to increase exercise engagement by older people. In the following quotes, 2 participants describe the importance of “seeing” another older person in web-based videos:

> It was great to see an older person do the exercises, really motivating to see someone my age doing the program. The videos were at the right pace, and I like how they got harder over time. It was fun. [P17]
> The exercise videos are motivating because I feel like I am doing it with someone—it’s interactive and fun. Following a book can be boring. [P5]

Many stakeholders commented positively on the video design that incorporated slow demonstration and simple dialog that aimed to teach aged care service users how to perform safe and effective home exercise:

> The physiotherapist in the videos demonstrated the exercises slowly and explained things easily. I was really surprised how the residents were able to follow everything without any help. [Residential coach, metro]
> Having online exercise resources really helps because people aren’t familiar with exercise techniques, they can follow their prescribed video and it helps keep their exercise dose up. [Home care physiotherapist, regional]

Many of the physiotherapists interviewed commented that the TOP UP program was complex and challenged participants to navigate different apps and printed resources such as exercise diaries while using Zoom. They suggested that simpler telephysiotherapy programs (or simpler mechanisms for accessing program components) could be developed to enhance the user experience and minimize program dropout. One physiotherapist commented:

> I think having an easy to navigate, no fuss system where our clients can look up an exercise, record their exercise program and any problems they may have had, a fall, etc. I think an app where physiotherapists could get access to this information easily during a session and to help prepare for another session would be useful. [Home care physiotherapist, rural]

Theme 4: A Flexible Reablement Approach Fosters Autonomy

TOP UP is designed to encourage older adults to take a lead in their program planning; flexibility is emphasized, including choices about what resources to use (printed and web-based) and what skills they wish to develop that would enable them to engage in activities they found most important:

> I liked how it started easily, and I moved my way up the program. There is structure to the program, and you commit to it. I often plan to do a session but if something comes up, I make an appointment with myself to make sure I do it another time. [P16]
> I liked that I could stop and start the videos according to my own needs on the day. [P1]

This can be described as a reablement approach, and the physiotherapist and coaches were encouraged to build the aged care service users’ physical capability and support them in transferring their new skills to access other activities in their community independently:

> Residents lack enough physical activity here, sometimes we are short staffed, and sometimes the staff don’t have time to help. It was great to see our residents on the TOP UP program improve their mobility and begin to walk to different activities on their own. [Residential manager, rural]

All stakeholders valued the reablement approach, and it was reported that TOP UP seemed to be a catalyst for reablement, as many of their clients began to engage in more socialization with friends and families and embrace other physical activities as they became stronger and more mobile:

> Physios and coaches can work together to ensure that the participant becomes independent and autonomous in their use of telehealth and do more exercise as the program progresses. As they improved, we had discussions with them and their coach about how they could do more outdoor walking. [Home care physiotherapist, rural]
> I was surprised about the other quality of life benefits of telehealth, talking to their physios on zoom, seeing their support workers in this new way, learning how to get out and about in the community, all seemed to reduce social isolation, which is so important for our customers. [Home care manager, rural]

Theme 5: Telephysiotherapy Can Be Safe, Effective, and Acceptable for Many

Most interviewees regarded TOP UP as a safe, effective, and acceptable program. Interviewees reported positive physical and quality of life improvements:

> I think it’s fabulous. I wouldn’t have imagined that I would be given the opportunity to get physio. Physically, I can walk further. My breathing is better. I’m stronger, it gives you more independence. [P1]
> Telehealth has not only helped my customer’s strength, mobility and coordination, but it seemed to help their overall quality of life, they seemed happier and more confident to walk. [Residential care manager, rural]

Many interviewees reported that the combination of physiotherapist-led instructional exercise videos and supervision by trained support workers increased the safety of the TOP UP:

> I think having a physio run exercises in the videos gives the intervention more authority, frees up my...
time to motivate the residents and keep them safe. [Residential coach, remote]

I think having the care worker there with the client to help set up Zoom, hold the iPad, and angling the video so I can see them clearly makes the program safer and more successful. [Home care physiotherapist, rural]

However, TOP UP was not considered to be suitable for all aged care service users. All stakeholders agreed that telehealth presents challenges for frail clients in residential aged care, who often have higher levels of mobility and sensory and cognitive disability. A total of 2 cognitive and sensory impaired aged care service users found using Zoom to “see” their physiotherapist frustrating and as a result, pulled out of the program:

First of all, not all dementia residents get used to it, and second, people with hearing and vision problems struggle to follow. [Residential aged care manager, rural]

Some physiotherapists would hesitate to use telehealth without local support for those aged care service users with high fall risks:

For people who are mostly independent I wasn’t worried, but if I did have someone who was who was frailer and there was no one there with them I was worried they might fall. [Home care physiotherapist, regional]

Some aged care service users and managers suggested that although telephysiotherapy is a good secondary option, they would still prefer in-person physiotherapy, especially for older adults with more complex needs:

I prefer a blend of face-to-face physio and telehealth. I need some hands-on physio from time to time to manage the arthritis in my back, but I liked the telehealth program because I could follow the physio exercise videos at home, it was so convenient. [P18]

It appears that a hybrid model that incorporates a blend of face-to-face physiotherapy and web-based exercise resources, such as exercise videos, was viewed as particularly acceptable for those with significant health challenges:

I don’t know if someone with severe dementia or disabilities would be able to access telehealth. I also think a lot of clients would like a hybrid telehealth model starting with a face-to-face assessment. [Home care physiotherapist, rural]

Finally, our screening process uncovered many technological hesitations and potential telehealth data concerns that prevented the recruitment of many potential aged care service users into the TOP UP trial:

There is some hesitancy around technology use due to recent cybersecurity anxiety in the community— for example the Optus and Medibank breaches. [Home care manager, rural]

Theme 6: Organizational Commitment Is Required to Embed Telephysiotherapy

Interviewees explained that considerable organizational commitment is required to embed telehealth programs such as TOP UP in aged care. Sufficient investment is required to train staff, conduct more meetings with their physiotherapy service providers to plan for the development of a new service, such as telephysiotherapy, prioritize TOP UP sessions within busy service schedules, and, where necessary, direct funds toward supportive technology. Some coaches and physiotherapists commented that the use of devices such as large iPads and smart televisions enhanced telehealth engagement by improving the visibility and hearing experience of service users:

Zoom worked well when we connected the iPad to the TV, we were able to turn the volume of the TV up so the resident could hear better. It also gives a bigger picture as well, so they can see the physio better. [Residential care coach, rural]

However, such equipment can be costly, and telehealth-specific funding was raised by physiotherapy, aged care managers, and coaches as a key condition for ongoing sustainability of telephysiotherapy in aged care:

I think that maybe there needs to be funding support. Telehealth is an important and easy way to increase access and uptake. One physio could service several homes in a full-time caseload. [Residential care physiotherapist, rural]

TOP UP required 3 people to be available for appointments (the older person, their coach, and the physiotherapist on Zoom); thus, scheduling was more challenging than 2-person face-to-face health care interactions:

There are always challenges whenever it comes to scheduling, especially during COVID when we were short of staff. But if you have a good relationship with your physiotherapy provider, who is responsive to time slot suggestions, then our scheduling team could work their magic and get it all booked. [Home care manager, rural]

Training was provided to older adults to increase their confidence using an iPad, our website, and relevant apps (Zoom and StandingTall). Coaches were trained to increase their level of comfort by navigating the TOP UP website and Zoom. Physiotherapists were trained to deliver effective telephysiotherapy assessments using Zoom and provided strategies to enhance relationship development with older adults and their coaches. All interviewees highlighted this training as an important factor in overcoming “telehealth hesitancy” both for service users and program providers:

There was a lot of telehealth hesitancy at the beginning, but with education they slowly got quite comfortable in doing it. [Home care physiotherapist, metro]

There is a need to have some general training so we [physiotherapists] know how to use it [telehealth technology] effectively: make sure your voice is coming through, how to pace instruction so our clients
understand us. The coaches and customers need training to know how to set up a shot, to make sure that they are visible to ensure that the client becomes independent and autonomous in their use of telehealth. [Home care physiotherapist, rural]

The aged care service managers also noted the challenge of training adequate numbers of care workers to facilitate TOP UP and ensure that the coaches are safe and competent:

There is a need to train a large proportion of our support workforce so that we have more trained staff who know how the program works, how to use technology and how to supervise our customers safely. [Home care manager, rural]

Some aged care service managers and physiotherapists indicated that more frequent and more detailed web-based exercise training programs would be useful to improve the skill level of a wider group of support staff:

It is very important to have lots of staff trained. For example, if the regular coach is sick, another staff

member could take over and keep the program going. [Home care manager, rural]

All stakeholders indicated the need for specific investment into better internet connectivity to ensure the sustainability of future telephysiotherapy programs:

I’ve found is there are still a lot of places in rural Australia where older people don’t have fast Internet, they don’t have smart TVs, or they don’t have the technology that metro places have. People are ready to engage with telehealth, but there’s no infrastructure in rural areas. [Residential physiotherapist, metro]

Implementation Guidance Through the Lens of the NASSS Framework

TOP UP appears to be well positioned for sustainable adoption, and learnings from this study have informed the translation of telephysiotherapy services by our aged care partners into practice. Table 1 uses the NASSS framework to help explain TOP UP’s successes and failures and explore the facilitators required to embed similar telephysiotherapy programs in aged care.
Table 1. An overview of TOP UP implementation guidance in relation to the Nonadoption, Abandonment and Challenges to the Scale-Up, Spread and Sustainability of Health and Care Technologies framework domains.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Definition of domain</th>
<th>Implementation guidance derived from study findings</th>
<th>Illustrative quotes from interviewees</th>
</tr>
</thead>
<tbody>
<tr>
<td>The condition</td>
<td>The suitability of the participant’s attributes/needs and their interaction with the intervention.</td>
<td>TOP UP is suitable for aged care service users with mobility challenges who can walk short distances. It is not suitable for those with significant sensory and cognitive disability.</td>
<td>“Someone with severe dementia or severe disabilities would not be able to have that skill to access telehealth.” [Residential coach, rural]</td>
</tr>
<tr>
<td>The technology</td>
<td>Technical features related to the usability of telehealth and its support requirements.</td>
<td>TOP UP requires access to the internet via an iPad or similar device. Aged care service users do not need technological skills due to the weekly support they received from trained care workers to help them use the iPad, navigate Zoom, and access exercise videos on a website. However, basic technological skills were often developed, which increased autonomy.</td>
<td>“One of my clients is really good with technology but other clients need my help to turn on the iPad and follow the program.” [Residential coach, metro]</td>
</tr>
<tr>
<td>The value proposition</td>
<td>The value proposition of telehealth for upstream end users (aged care service providers) and downstream users (physiotherapists and their clients).</td>
<td>All stakeholders saw telephysiotherapy as a valuable addition because of its convenience and perceived effectiveness, especially for those with poor mobility or who are living in rural or remote areas. The value proposition for telehealth to treat musculoskeletal pain is less as stakeholders prefer a more “hands-on” experience. A hybrid model would add value for some.</td>
<td>“Telehealth would save us time and travel and help us to see more people.” [Home care physiotherapist, regional]</td>
</tr>
<tr>
<td>The adopter system</td>
<td>The ongoing investment required to support the telehealth intervention and the ongoing acceptability of stakeholders.</td>
<td>TOP UP requires consistent investment in training, human (physio, coaches), and physical infrastructure (devices, fast internet, senior-friendly exercise resources) to create sustainable success. However, high levels of system support are likely to be reinforced as positive returns on investment due to their perceived positive impacts on mobility and well-being.</td>
<td>“TOP UP is more than just a fall prevention program, it offers a truly re-ablement focus where our clients can build their strength and balance and get out into the community again. I think many of our clients could benefit from telehealth.” [Home care manager, rural]</td>
</tr>
<tr>
<td>The organization</td>
<td>An organization’s capacity to embrace the telehealth intervention and the supports required to establish and maintain it as a viable service offering.</td>
<td>Not all aged care services chose to participate in TOP UP because of the perceived burden of working with technology. Providers who joined TOP UP wanted to investigate telehealth’s impact on access to fall prevention and mobility programs, in areas where there are physiotherapy shortages. Providers offered considerable support via technology provision, extra administration support for scheduling of telephysiotherapy sessions, and enough care workers to support the program.</td>
<td>“I was surprised at how easy telehealth was to get started. We gave the clients an iPad and the assistance the care workers gave them was important to help them engage with telehealth. Our scheduling team are fantastic, and they managed to solve the scheduling challenges really well.” [Home care manager, rural]</td>
</tr>
<tr>
<td>The wider context</td>
<td>The wider organizational and policy impacts on telehealth uptake and sustainability.</td>
<td>Stakeholders agreed that funders need to provide telehealth-specific funding and education for interventions such as TOP UP to reduce technology hesitation and improve telehealth systems that enhance its adoption and sustainability.</td>
<td>“I feel that people would be greatly advantaged if there was a separate pocket of funding for allied health so that we could afford to deliver ongoing telehealth” [Home care manager, rural]</td>
</tr>
</tbody>
</table>

Discussion

Principal Findings

This study, which included participants such as older adults, physiotherapists, aged care support workers, and managers in the TOP UP trial, offers valuable insights. Our thematic analysis identified key factors for the telephysiotherapy program’s acceptability, including advice from physiotherapists, consistent support from trained care workers, older adult-friendly web-based exercise resources, and a flexible reablement approach. The interview data supported multiple themes, suggesting that the synergistic integration of these ingredients within the TOP UP contributed to its high acceptability. The discussion explores the impact of single components and emphasizes their combined contribution to TOP UP’s acceptability.

TOP UP Study Is Acceptable

Acceptability is an important consideration in the design and implementation of complex health care interventions, such as TOP UP [36]. Our findings align with the increasing body of literature indicating acceptance of telehealth among older adults in community settings despite high levels of technology hesitation [37,38]. A cohort study of a telehealth program incorporating physiotherapy for rural older adults found that telehealth was safe (no adverse events) and feasible (average telehealth attendance 85%) [38]. A 2021 cross-sectional survey of health care providers further affirmed increasing telehealth acceptability over time among homebound older adults [39].

Barriers and Facilitators Related to Telehealth Adoption

TOP UP identified several barriers and facilitators that enabled aged care service users to overcome high levels of technology
hesitation and, if appropriately addressed, could improve the translation of telehealth programs into aged care [40]. TOP UP’s qualitative findings are reflected in the literature, which demonstrate that barriers related to this population’s innate technology hesitation and greater sensory, physical, and cognitive impairments could be addressed by the provision of local support, internet-connected devices, fast internet, and appropriate telehealth training can mitigate these barriers [39,41,42].

A recent US survey of physician providers of homebound older adults during the COVID-19 pandemic revealed that a significant proportion of their patients were technology avoidant [39] (only a third of their patients had used video-based telehealth before, 310/873, 35.5%). Among patients who had not used telehealth before, providers deemed that one in 4 (153/563, 27.5%) of their patients would not be able “interact over video” due to cognitive or sensory impairments. This survey found other barriers: providers lacked knowledge of their patients’ internet connectivity, and participants faced financial constraints in obtaining internet plans and were unable to pay for internet plans or video-capable devices. Similar findings emerged in the TOP UP, where most trial participants had limited access (10/18, 56%) to video-capable devices, limited telehealth experience (4/18, 22%), and low telehealth readiness (Technology Readiness Index 2 out of 5). Addressing barriers related to the purchase of telehealth infrastructure and providing local support can facilitate wider acceptance within aged care settings.

A recent qualitative exploration of factors influencing acceptability in dementia management revealed that videoconferencing had potential benefits over in-person appointments by improving access to care for those with mobility limitations and reducing the stress associated with clinic appointments [43]. A crucial insight from this study emphasized the necessity of technical support and telehealth training involving information on how to access and use different telehealth apps and tips for setting up the video camera for maximum visibility. Similarly, another study examining telehealth’s role in enhancing oncology care for older adults emphasized that appropriate technology training integrated into the screening process and program delivery could enhance telehealth adoption [44]. These studies align with TOP UP’s findings that emphasized the delivery of appropriate education at screening and recruitment to reduce technology avoidant behaviors, preprogram technology training to support adoption, and training to troubleshoot any emerging technology issues to enhance sustainability.

TOP UP demonstrated that behavior change training for physiotherapists and coaches in health coaching techniques, motivational interviewing, and collaborative goal setting can facilitate telehealth adoption. Behavior change training has been shown to increase therapeutic alliance and enhance exercise program outcomes in other studies [45]. A strong therapeutic alliance has been identified as a crucial facilitator in previous telehealth interventions [46]. In our study, physiotherapists, coaches, and aged care service users found telehealth suitable for effective behavioral change coaching and suggested that specific training on skills to enhance therapeutic alliance is important to augment telehealth acceptability. Specific examples included targeted training on using Zoom emojis to acknowledge client achievements and building a personal connection through virtual tours of the older person’s home and garden. However, they noted that establishing a successful therapeutic alliance through telehealth demands more time compared with in-person sessions, potentially increasing program costs.

**Telehealth Can Provide Key Ingredients for Behavior Change**

TOP UP was co-designed to incorporate the COM-B model to create positive behavior changes related to exercise adherence [17]. Recent data from the Australian Institute for Health and Welfare have shown the critical significance of addressing insufficient physical activity in older individuals, given their 50% contribution to 2.5% of the overall disease burden in Australia [47]. Consequently, increasing motivation and opportunities for exercise in this demographic is crucial in mitigating the adverse health consequences stemming from sedentary behavior [19] and in supporting the efficiency of the health care system [48]. TOP UP’s tailored approach and use of older adult–friendly resources appeared to increase the capability (C) of older adults to exercise. The program provided increased opportunities (O) for exercise by facilitating increased access to physiotherapists. Furthermore, TOP UP heightened motivation (M) through its reabatement approach, goal-setting mechanisms, and cultivation of enjoyment via companionable coaching [16].

The TOP UP program strategically incorporated the principles of SDT to promote increased exercise adherence. According to SDT, intrinsic motivation thrives when individuals perceive a sense of autonomy and control over their activities [49]. Our study findings suggest that the aged care service users valued the opportunity to regain independence through self-directed exercise. The TOP UP program effectively nurtured feelings of competence through its personalized and progressive exercise routines program, fostering a sense of relatedness through local support and the rapport established during the telephysiotherapy sessions that actively promoted enjoyment. This observation aligns with the systematic review by Teixeira et al [50] on SDT and exercise adherence, affirming the positive correlation between intrinsic motivation, enjoyment, personal achievement, and heightened program acceptability.

Our study has provided insights into the potential explanatory effects of the social learning theory by Bandura [51] and Motivational Theory of Role Modeling in supporting the high acceptability of TOP UP. The social learning theory by Bandura [51] underscores the significance of observation and imitation in driving behavior change. When individuals perceive the modeled behavior as valuable, and the model possesses an admired status while being relatable, the likelihood of behavioral change increases. In this context, physiotherapists, esteemed as exercise professionals in the community [52], played a crucial role in enhancing the perceived value of the TOP UP program. Furthermore, the Motivational Theory of Role Modeling highlights another critical aspect of TOP UP’s acceptability [53]. Many interviewees emphasized the importance of including older adults as role models in exercise videos. Both theories
suggest that the inclusion of older role models was a pivotal factor inspiring behavioral change, explaining the positive reception of TOP UP exercise videos.

**Scale-Up and Sustainability of Telephysiotherapy in Aged Care**

Telehealth has emerged as a prominent method for implementing scalable health care interventions, a trend that has intensified during the COVID-19 pandemic [54]. However, the challenge of sustaining these programs is pressing, as is evident from reports of high participant attrition rates in telehealth-led exercise programs [55]. Successfully delivering cost-effective exercise programs to frail older adults with multiple comorbidities in the aged care environment is challenging and complex, demanding significant resourcing [8]. Insights gained from the NASSS framework [28] underscore the need for careful screening of older adults for telephysiotherapy participation and the provision of targeted training to all stakeholders to enhance its feasibility. Our analysis indicates that while TOP UP was acceptable, a hybrid model of virtual care that combines in-person initial assessments, subsequent synchronous telephysiotherapy sessions for program progression, and the integration of local support and older adult–friendly web-based exercise resources may further increase telephysiotherapy uptake and sustainability in aged care.

Although the cost-effectiveness analysis of TOP UP is pending, our qualitative observations indicate that establishing telephysiotherapy programs requires substantial investment in both physical and human infrastructure. The telehealth literature discusses the critical role governments play in developing policies and guidelines to foster telehealth adoption [56]. Our interviews revealed a consensus on the need for dedicated funding for telehealth to enhance adoption and sustainability.

**Strengths, Limitations, and Future Studies**

This qualitative study had several strengths. It triangulates empirical data relating to the uptake and sustainability of telephysiotherapy in aged care from 4 perspectives: older adults receiving physiotherapy within aged care services, physiotherapists, trained support workers who deliver the intervention, and aged care managers who are charged with case management and overseeing aged care service resource allocation and delivery. Our partnerships with aged care providers and their ongoing input in the research have enabled us to develop a deep understanding of how the TOP UP program was delivered in aged care and, if proven effective, this will speed up its translation into wider practice [14].

Qualitative research serves as a valuable tool for refining program design, deepening insights into the outcomes of quantitative research, and offering valuable guidance for enhancing the implementation of complex interventions such as telephysiotherapy in aged care [26]. In this study, we adopted a broad sampling strategy aimed at delivering a rich description of diverse intervention experiences, enlisting the perspectives of 18 older adults encompassing a range of sociodemographic characteristics distributed across 10 distinct sites. Moreover, the inclusion of independent physiotherapists, separate from both the aged care service partners and the research team, in our study design may have reduced potential social desirability bias, enhancing the credibility of our findings [30].

Several limitations of this study necessitate careful consideration. TOP UP excluded participants from culturally and linguistically diverse backgrounds and thus presented a notable gap in our understanding of their experiences. To address this gap, future trials that prioritize the inclusion of culturally and linguistically diverse communities are required. In addition, although the interviewee cohort was purposefully selected to encompass maximum variation, it is essential to acknowledge that this pool primarily consisted of individuals who voluntarily participated in the trial, potentially predisposing them to higher levels of exercise engagement and receptiveness to telehealth. Consequently, this may limit the generalizability of our findings.

Several aged care service users and coaches were interviewed by either a physiotherapist or aged care service provider who delivered the program. This can lead to social desirability biases, which may undermine the credibility of the study results [57]. Given this context, aged care service users and staff might hesitate to openly share negative experiences with their interviewers despite the research team’s assurances that their feedback would have no bearing on their ongoing care or employment status. To mitigate this bias, interviewers made concerted efforts to positioning themselves as eager learners, actively encouraging interviewees to share their “insider” perspectives on quality improvement and expressing genuine appreciation for any criticism offered. Future larger-scale mixed methods studies should be designed to enhance research quality and further explore the impact of telephysiotherapy uptake and sustainability in aged care while carefully addressing social desirability bias.

This study suggests a need for the development of simplified telephysiotherapy exercise programs to facilitate greater adoption in aged care. A recent scoping review conducted in 2021, examining the barriers and facilitators to the use of telehealth by older adults, found several impediments associated with current technology, including challenges related to small screen, text size, small icons, insufficient color contrast between text and background, and complex functionality [58]. The review also identified ease of use as a key facilitator of telehealth adoption. Some TOP UP stakeholders interviewed indicated a preference for simplified functionality tailored to this demographic. Respondents expressed a desire for telehealth programs that incorporated TOP UP program features such as Zoom, exercise diaries, and videos, into one user-friendly application. These findings advocate for further research aimed at enhancing the user experience.

**Conclusions**

This interview study explored the program experiences of aged care service users, physiotherapists, and aged care staff involved in the TOP UP trial, a telehealth-led exercise program designed to improve mobility, reduce falls, and enhance quality of life. All stakeholders indicated high program acceptability, underscored by its safety, and perceived effectiveness. The thematic analysis uncovered key insights: TOP UP’s provision of convenient access to physiotherapy services for aged care

https://aging.jmir.org/2024/1/e53010
recipients, the positive impact of tailored physiotherapy coupled with local support on exercise motivation, the effectiveness of engaging older adult–friendly resources in fostering program adherence, and the facilitation of greater independence through a flexible reablement approach. This study emphasizes the importance of sustained organizational commitment for the successful implementation of telephysiotherapy programs, such as TOP UP, highlighting the need for training and external funding to ensure telephysiotherapy’s adoption and sustainability.

Acknowledgments
RD receives scholarship funding and JO receives salary support from the National Health and Medical Research Council–funded Centre for Research Excellence—Prevention of Fall-related Injuries. MP holds a National Health and Medical Research Council of Australia Early Career Fellowship. KN is employed by Whiddon and NB is employed by Uniting AgeWell. The TOP UP study was funded by Dementia Australia and Aged Care Research and Industry Innovation Australia. The funders and organizations had no role in the trial design and will not have any role during its execution, analyses, interpretation of the data, or decision to submit the results.

Data Availability
The data sets generated during this study are not publicly available because of the confidentiality promised to the participants as part of the informed consent process. However, data sets are available from the corresponding author upon reasonable request.

Authors' Contributions
All authors reviewed the thematic analysis. RD, AH, and KN drafted the semistructured interview guide. RD drafted the manuscript as the lead author. All authors critically revised the manuscript for intellectual content.

Conflicts of Interest
RD receives scholarship funding, and JO receives salary support from the National Health and Medical Research Council–funded Centre for Research Excellence—Prevention of Fall-related Injuries. MP holds a National Health and Medical Research Council of Australia Early Career Fellowship. KN is employed by Whiddon, and NB is employed by Uniting AgeWell. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Semistructured interview guide. [DOCX File, 17 KB - aging_v71e53010_app1.docx]

Multimedia Appendix 2
The Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist. [DOCX File, 16 KB - aging_v71e53010_app2.docx]

References


27. Patton MQ. Enhancing the quality and credibility of qualitative analysis. Health Serv Res 1999 Dec;34(5 Pt 2):1189-1208 [FREE Full text] [Medline: 10591279]


Abbreviations

**COM-B**: capability, opportunity, motivation, and behavior  
**NASSS**: Nonadoption, Abandonment and Challenges to the Scale-Up, Spread and Sustainability of Health and Care Technologies  
**SDT**: Self-Determination Theory
Acceptance of a Digital Assistant (Anne4Care) for Older Adult Immigrants Living With Dementia: Qualitative Descriptive Study

Marloes Bults¹, PhD; Catharina Margaretha van Leersum²,³, PhD; Theodorus Johannes Josef Olthuis⁴, MA; Egbert Siebrand⁵, MSc; Zohrah Malik¹, MSc; Lili Liu⁶, PhD; Antonio Miguel-Cruz⁶,⁷,⁸, DSC; Jan Seerp Jukema⁹, PhD; Marjolein Elisabeth Maria den Ouden¹,⁴, PhD

¹Technology, Health & Care Research Group, Saxion University of Applied Sciences, Enschede, Netherlands
²Department of Technology, Policy, and Society, Faculty of Behavioural, Management and Social Sciences, University of Twente, Enschede, Netherlands
³Faculty of Humanities, Open University, Heerlen, Netherlands
⁴Care & Technology Research Group, Regional Community College of Twente, Hengelo, Netherlands
⁵Research Group Ethics and Technology, Saxion University of Applied Sciences, Enschede, Netherlands
⁶School of Public Health Sciences, Faculty of Health, University of Waterloo, Waterloo, ON, Canada
⁷Department of Occupational Therapy, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada
⁸GRRIT Hub Glenrose Rehabilitation Research, Innovation & Technology, Glenrose Rehabilitation Hospital, Edmonton, AB, Canada
⁹Smart Health Research Group, Saxion University of Applied Sciences, Enschede, Netherlands

Corresponding Author:
Marloes Bults, PhD
Technology, Health & Care Research Group
Saxion University of Applied Sciences
M H Tromplaan 28 Enschede
Enschede, 7513 AB
Netherlands
Phone: 31 620946213
Email: m.bults@saxion.nl

Abstract

Background: There is a need to develop and coordinate dementia care plans that use assistive technology for vulnerable groups such as immigrant populations. However, immigrant populations are seldom included in various stages of the development and implementation of assistive technology, which does not optimize technology acceptance.

Objective: This study aims to gain an in-depth understanding of the acceptance of a digital personal assistant, called Anne4Care, by older adult immigrants living with dementia in their own homes.

Methods: This study used a qualitative descriptive research design with naturalistic inquiry. A total of 13 older adults participated in this study. The participants were invited for 2 interviews. After an introduction of Anne4Care, the first interview examined the lives and needs of participants, their expectations, and previous experiences with assistive technology in daily life. Four months later, the second interview sought to understand facilitators and barriers, suggestions for modifications, and the role of health care professionals. Three semistructured interviews were conducted with health care professionals to examine the roles and challenges they experienced in the use and implementation of Anne4Care. Content analysis, using NVivo11, was performed on all transcripts.

Results: All 13 participants had an immigration background. There were 10 male and 3 female participants, with ages ranging from 52 to 83 years. Participants were diagnosed with an early-stage form of dementia or acquired brain injury. None of the older adult participants knew or used digital assistive technology at the beginning. They obtained assistance from health care professionals and family caregivers who explained and set up the technology. Four themes were found to be critical aspects of the acceptance of the digital personal assistant Anne4Care: (1) use of Anne4Care, (2) positive aspects of Anne4Care, (3) challenges with Anne4Care, and (4) expectations. Assistance at first increased the burden on health care professionals and families. After the initial effort, most health care professionals and families experienced that Anne4Care reduced their tasks and stress. Contributions of Anne4Care included companionship, help with daily tasks, and opportunities to communicate in multiple languages. On the other hand, some participants expressed anxiety toward the use of Anne4Care. Furthermore, the platform required an internet connection at home and Anne4Care could not be used outside the home.
Conclusions: Although older adult immigrants living with dementia had no previous experience with digital assistive technology specifically, the acceptance of the digital personal assistant, called Anne4Care, by older adult immigrants living with dementia was rather high. The digital assistant can be further developed to allow for interactive conversations and for use outside of one’s home. Participation of end users during various stages of the development, refinement, and implementation of health technology innovations is of utmost importance to maximize technology acceptance.

(JMIR Aging 2024;7:e50219) doi:10.2196/50219

KEYWORDS
assistive technology; technology acceptance; immigrant; dementia; marginalized older adults

Introduction

Background
Dementia is a global health problem associated with emotional and financial challenges for people living with dementia, their relatives, health care professionals, and health organizations [1,2]. Worldwide, approximately 47 million people are diagnosed with dementia [3]. Around 280,000 persons are presently living with dementia in the Netherlands among a population of nearly 19 million [4]. The number of people living with dementia is expected to increase in the coming years. People with an immigration background are at an increased risk of developing dementia and are more likely to experience barriers in accessing dementia services and care, which may lead to health inequities and a reduction of quality of life [5-7]. Hence, there is a need to develop and coordinate dementia care plans, including the use of technology, for marginalized older adults living with dementia, such as those with an immigration background and low literacy levels [8].

Health Literacy and Health Technology
Health literacy has been defined as “the degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions” [9]. Adequate health literacy and access to services among immigrant populations are important to empower, support or facilitate these persons to become active participants in their health [10]. Persons with limited health literacy have difficulty finding, understanding, and applying information about health and health care. Health technology–based solutions that support health promotion, for example, mental and physical health, offer opportunities to increase health literacy in vulnerable populations [11,12].

The World Health Organization reports that innovative health technologies are promising tools to enhance knowledge, skills, and coping mechanisms to improve the daily lives of people with dementia and their caregivers [13]. In recent years, a variety of technologies to support self-management have been developed for people with dementia and their caregivers. Studies have focused on the effect of these technologies on the physical, mental, and social well-being of people with dementia and their relatives and families [14,15]. Technology seems to stimulate cognitive function and communication skills and reduce loneliness among people with dementia, but these results are personal and context-dependent [14,15]. Acceptance of technology is a major challenge and prerequisite for the implementation of technology in practice. For successful technology implementation in the daily lives of people, it is important to have insight into the acceptance of technology among end users. Although studies have examined the acceptance of technology among people living with dementia and their caregivers [16-18], the number of studies that focus specifically on older adult immigrants living with dementia is minimal.

Citizen Science
Co-design and the involvement of people with dementia and their professional and family caregivers are crucial elements for the acceptance of technology [17]. However, immigrant populations are seldom included in stages of development, refinement, and implementation of assistive technology innovations [19]. This can be associated with low levels of acceptance of health technology innovations, which can contribute to health inequities among populations. One opportunity to increase the involvement of end users is through citizen science, which is an approach that engages end users to be partners in research so that their experiences and needs are considered [20]. Citizen science is a powerful approach to include public participation in research as well as optimize acceptance of technologies [20,21]. In citizen science, scientific principles and methods are used by nonprofessional “scientists” in close collaboration with scientific researchers [22]. The involvement of older adults in the development, refinement, and implementation of technology, acknowledging their expertise and needs, and working together in short iterations to adapt the technology for their specific needs are reported to be valuable elements by scientific researchers, older adults, and health care professionals [23].

Aim
This qualitative descriptive study aimed to provide an in-depth understanding of the acceptance of a digital personal assistant, called Anne4Care, among older adult immigrants living with dementia by using a citizen science approach.

Methods

Research Design
This study used a qualitative descriptive research design with a naturalistic inquiry [24] and citizen science principles [20]. In this study, Anne4Care was extended with the development of a Turkish version. All the involved older adults had the task of testing the device as well as communicating with each other, their care professionals, and the Anne4Care help desk. For example, there were workshop-like meetings in which all shared
experiences, assisted each other with difficulties, and expressed the need to further develop the Turkish version of Anne4Care. Furthermore, participants actively collaborated with scientific researchers, for example, in discussing the topics for the interview guide and analyzing the data from the interviews.

Aligning with citizen science principles there was a close collaboration between older adults participating as co-researchers, health care professionals, and scientific researchers. A detailed description of how older adults with an immigration background were engaged in this study is described in a separate paper [23]. The team of researchers closely collaborated with a group of citizens representing the target group and health care professionals in their role as co-researcher, during all phases of this study: recruitment, obtaining informed consent, data collection, and analysis.

Setting
This research was part of the TOPFIT Citizenlab program, a research and innovation program in which citizens, health care professionals, and companies join forces with scientific researchers to develop and implement technology for health.

The Technology
Anne4Care is a digital personal assistant that includes video-calling, a personal agenda, medication reminders, reading the news, and games that can be used in one’s home. Anne4Care is a technological platform created to help people with dementia to continue living independently in their own homes and supporting caregivers in their tasks [25]. Anne4Care included hardware as well as a software platform. Anne4Care is available in Dutch, German, Italian, and English. The company was developing a Turkish version of Anne4Care and saw the embedding of Anne4Care in homes of older adults with an immigration background as an opportunity to test, improve, and implement the latest version.

Recruitment Strategy
The recruitment of participants was performed by 3 health care professionals from 2 health care organizations. These organizations provide care for clients with an immigration background with cognitive impairments.

Inclusion criteria were as follows:
- Diagnosed with an early-stage form of dementia or acquired brain injury;
- Having an immigrant background;
- Visiting the activity program of 1 of the 2 participating health care organizations

Exclusion criteria were as follows:
- Diagnosed with a severe stage form of dementia limiting their ability to participate

All clients of the 2 health care organizations (IMEAN Consultancy & Care and Alifa Wellbeing Older Adults) were invited by their health care professionals to participate in this study. All older adult participants had an immigration background, that is, 1 came from Britain and the other 12 came from Turkey (Table 1). They visited the activity program of 1 of the 2 health care organizations, which are situated in the Twente region of the Netherlands. In addition to an immigration background, all 13 older adult participants, or co-researchers, had an early-stage form of dementia or acquired brain injury. The sample included 10 male and 3 female participants, and their ages ranged from 52 to 83 years. There were no exclusion criteria based on digital literacy, that is, participants did not need any experience with technology. Internet access was provided to participants who did not have an internet connection at home.

Table 1. Demographic characteristics of the participants (N=13).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td></td>
<td>Range</td>
</tr>
<tr>
<td>Age group (years), n (%)</td>
<td>45-54</td>
</tr>
<tr>
<td></td>
<td>55-64</td>
</tr>
<tr>
<td></td>
<td>65-74</td>
</tr>
<tr>
<td></td>
<td>75-84</td>
</tr>
<tr>
<td>Sex at birth, n (%)</td>
<td>Female</td>
</tr>
<tr>
<td></td>
<td>Male</td>
</tr>
<tr>
<td>Nationality, n (%)</td>
<td>Turkish</td>
</tr>
<tr>
<td></td>
<td>British</td>
</tr>
</tbody>
</table>
All 3 health care professionals involved in this study were female. They introduced Anne4Care to the older adults, assisted them, and were in close contact with the participants during all phases of this study. The health care professionals took care of the clients and spoke their native language, which created a safe and trustful environment. The health care professionals also served as a voice for the participants who could not communicate in the Dutch language.

Data Collection
Anne4Care was introduced to allow health care professionals to communicate with and monitor older adults. Data were collected between September 2020 and November 2021. Semistructured in-depth interviews took place at the care organization or at the home of the older adult immigrants. The location was chosen based on a participant’s preference. During the interviews, a care professional was present and served as an interpreter. Participants were invited for 2 interviews. The first interview was planned shortly after the introduction of Anne4Care in their home, and a second interview 4 months later. Participants had the opportunity to continue using Anne4Care after the data collection period, funded through a stimulation subsidy for eHealth at home during the COVID-19 pandemic. Five researchers with mixed credentials, training, occupation, location, and gender conducted the interviews (MB, CMvL, TJJO, ES, and ZM). To secure interrater reliability, the researchers had biweekly meetings to discuss the procedure and previous interviews. Two researchers were present during each interview; the teams of 2 were different for each interview. One was the main interviewer and the other took notes and asked additional questions. The follow-up interviews were conducted by the same researchers to ensure the established relationship between researchers and older adults.

The topic guide was developed in collaboration with 1 participant from the daycare facility and 3 health care professionals. Thirteen first interviews were conducted focusing on understanding the lives and care needs of the participants, learning about their expectations regarding Anne4Care, and sharing previous experiences with care and technologies, what facilitators and barriers they encountered while using technology (Multimedia Appendix 1). The first interviews lasted between 30 and 60 minutes. A total of 8-second interviews were conducted with the same participants who participated in the first interview (5 of the participants were not able to participate in the second interview due to COVID-19 illness). The second interview focused on the facilitators and barriers participants experienced when using Anne4Care, suggestions for modifications of Anne4Care, and the role of health care professionals in using technology (Multimedia Appendix 1). The second interview lasted between 20 and 45 minutes.

Three semistructured interviews were conducted with the 3 health care professionals. These interviews took place at the care organization and lasted for 60 minutes. The aim of these interviews was to talk about their role and the challenges they experienced as professionals in the use and implementation of Anne4Care (Multimedia Appendix 2). Furthermore, the findings of the interviews with the older adults were shared with the care professionals.

Data Analysis
All interviews were audio recorded and transcribed verbatim. The transcripts were made in English and Dutch. All Turkish spoken words were translated by an interpreter during the interviews. Only these translations were part of the transcripts. Content analysis of transcripts used an inductive approach [26,27]. We used the software package NVivo11 to support data coding. Open coding was used to identify relevant themes, there were no themes in advance of the data analysis. Several steps were taken to develop a code book. First, 3 researchers (MB, CMvL, and TJJO) performed the analysis of 2 transcripts and compared codes. A preliminary codebook was developed comprising these themes. Second, the coding of one transcript was discussed together with one older adult. This participant was motivated and had some previous experience with research and data analysis. Together with the co-researcher, additional themes were added to the code book. Last, the other transcripts were analyzed by 1 researcher (CMvL). The data analysis and application of codes were discussed during biweekly meetings with the research team. Data saturation was reached after analyzing the data obtained with all involved older adults and care professionals. During the data analysis, similar and confirmation of all findings appeared when analyzing and coding the transcripts. In preparation for the paper, the quotes in the raw data were translated into English.

Ethical Considerations
Ethics approval was obtained from the ethical advice committee of the University of Applied Sciences Saxion (reference number SEAC-2020-005). The participants were informed about the study before the start of the research period with Anne4Care. Thirteen participants gave written consent and were informed about their right to withdraw at any time. Data were anonymized, confidentiality was maintained, and the data will be retained for a period of 10 years after which they will be destroyed.

Results
Overview
The analysis of data revealed that the personal situation and perceptions of participants regarding access to dementia services and care (in short: care) were 2 underlying themes that described the adopter system from the older adult immigrants living with dementia. The code tree is presented in Multimedia Appendix 3.

Personal Situation
Personal situation considers the perceptions of older adult immigrants regarding their health status (physically and mentally), level of spirituality, quality of life, social and societal participation, and daily functioning. The personal lives of participants were diverse, with their immigration background as a main commonality. They were all born in another country and moved to the Netherlands for their work. Most of the participants worked in the textile or metal industry. The participants were all retired and visited the facility for daytime activities once or twice a week. On other days, their daily activities consisted of grocery shopping, housekeeping, or just...
doing nothing. Spirituality (eg, religion) was an important part of the lives of most participants. The physical and mental health of all participants is deteriorating, with illnesses that range in severity. In addition to dementia or acquired brain injury, some have health conditions such as diabetes, high blood pressure, and cataracts. These deteriorating health conditions had a major impact on their daily functioning as well as their quality of life. In the earlier years of their lives, some participants experienced challenges in participating in Dutch society, but this became even more difficult during the COVID-19 pandemic. Social relations were mainly with family members, and some had close contact with their neighbors.

Since I came here in the Netherlands, in the year 1977, I bought a house and since then until now I live in the same street, the same neighborhood with the same neighbors and this was always a very good network. It is a community with the Turkish and Dutch neighbors, in which I am a beloved man, I go to the mosque a lot and had several board positions for a while, such as a board member of the mosque. [Participant H]

Care
Care refers to the perceptions of participants regarding access to dementia services and care in including all assistive and care needs of the participants. All participants acknowledged the fact that they became older and had increasing trouble taking care of themselves. Some participants had troublesome experiences with care in the past or could not find suitable and personalized care. In most cases, family members, partners, children, or neighbors assisted and provided support and care. However, most participants perceived this as a burden for their relatives and, therefore, was not a desirable situation.

Now there is someone who assists in housekeeping, and our children assist with the more administrative tasks. Next-door there are some younger neighbors who offered to help for example with the garden. I have a lot of help from all of them. However, in the past I had a lot of frustrations with health care professionals. When we arranged a time, they did not show up and none made any record of my needs. Then the agency went bankrupt and the clients were left in the dark. [Participant G]

Acceptance of Anne4Care
In addition to the 2 underlying themes, the analysis of the acceptance of Anne4Care by older adult immigrants revealed four themes: (1) use of Anne4Care, (2) positive aspects of Anne4Care, (3) challenges with Anne4Care, and (4) expectations. In this section, results are presented for each theme. The code tree is presented in Multimedia Appendix 3. Multimedia Appendix 4 provides visual information about Anne4Care.

Use of Anne4Care
This theme refers to the acceptance and actual use of the digital personal assistant Anne4Care by older adult immigrants living with dementia in their own homes. The use of Anne4Care applies to how someone uses Anne4Care, what someone does with Anne4Care, and how health care professionals or families are involved. None of the participants knew or used an assistive technology similar to Anne4Care. All were unfamiliar with the existence of these types of digital assistive technologies. One participant acknowledged that technology like Anne4Care could assist health care professionals. In addition, the health care professionals underlined the potential of Anne4Care for supporting clients at a distance. Although the participants were unfamiliar with technologies like Anne4Care, they were familiar with technologies such as a doorbell with a flashlight or a talking clock, and 1 participant owned a robot vacuum cleaner.

I was so pleased when I saw the result, this robot really cleans everything. You can just leave your home and it will clean everywhere. [Participant L]

Anne4Care was mainly used as a memory assistive tool, for appointments and medication.

It is very useful. Anne4Care tells me when I must take my medication. She helps me to remember, she is tough and fun. I am very happy with it. [Participant M]

In addition to the agenda function and medication reminders, the game, radio, and newspaper functions of Anne4Care were used or requested by some participants. A health care professional would need to be aware of participants’ requests in order to activate the radio or newspaper functions.

He would like to receive more radio channels. He has a Turkish music channel, but would like to receive Dutch channels as well. We can add these channels easily to the list, I will ask him at a later moment which he would prefer. [Health care professional B]

Thus, the health care professionals and sometimes family were responsible for adding new functions as well as appointments in the agenda, and changes in the medication list. In the beginning, this costed time and was a source of burden. Some assistance to understand Anne4Care was needed at the start. However, after this initial adjustment, most health care professionals and families experienced the tablet as a task relieving as well as stress relieving. Anne4Care gave the reminders so that care partners did not have to keep track of everything during the day. Some participants admitted that continuous reminders from their partners made them angry, but reminders from Anne4Care were received more positively, causing fewer troubling situations at home. Furthermore, some participants preferred to update the agenda themselves if they could learn to work with Anne4Care. This feature is currently not possible with the platform.

I just need a keyboard to add appointments in my agenda. It is important for me to do this myself without any assistance, just some explanation and exercises in the beginning. It would be great if that would work! [Participant E]

After a few months, 3 participants decided to stop using Anne4Care. Two participants did not see the additional value, Anne4Care did not give them any new tools, and 1 participant stopped using them due to illness.
Positive Aspects of Anne4Care

This theme refers to participant experiences regarding the advantages and benefits of Anne4Care. During the interviews, participants were asked to share the positive aspects of Anne4Care. The avatar of Anne4Care was received positively by the participants. Coincidentally, “Anne” is also the Turkish word for mother. Although Anne does not look like a Turkish mom, she gave a feeling that there was someone in their homes because she talks, makes movements, and looks like a nurturing health care professional. Another positive aspect was the choice of language. Most participants chose their native language because Dutch was progressively more difficult to use with age and since the onset of dementia. However, some participants chose the Dutch language intentionally in order to develop and maintain their Dutch language skills.

With regard to positive aspects, the participants commented on the functionalities of Anne4Care: video calling, agenda, medication reminders, games, radio, and newspaper. The most positive aspect of the video calling was the quality and the size of the screen (respectively 10-12.3 inch diagonal). They could see the other person more clearly on the tablet than, for example, on their smartphone. The most positive aspect about Anne4Care was the agenda function with the reminders of appointments and the medication reminder function. These reminders were very essential for the participants’ personal life and health. One health care professional explained that any event or task could be added.

Take for example the timing for their regular prayers. These are essential for someone’s life, and we can easily add these into the agenda. [Health care professional]

Another participant talked about the assistance of Anne4Care in the daily cooking routine.

There are reminders when I need to start cooking, but also already before which groceries I have to purchase. After a while Anne4Care asks: ‘did you turn off the stove?’ That is very helpful and important for me. [Participant K]

The game function was experienced as a fun activity to do during the day. Furthermore, the radio and newspapers provided by Anne4Care were perceived to be valuable. For example, some radio channels with music from their past gave the participants an opportunity to escape from their current time and place. The newspapers were seen as essential to keep up to date with current events; the read-aloud option was an asset.

It is wonderful that the radio and newspapers can provide the news into my home. The news keeps me up to date, and I also know what happens in Turkey where my family is. [Participant H]

Challenges With Anne4Care

After using Anne4Care, the participants were asked about any challenges they experienced. As mentioned, the agenda and medication reminder functions were experienced as positive. However, the video calling option presented challenges. The main challenge was in making a connection with others. A video call required both the caller and receiver to activate Anne4Care. Therefore, all participants had to first send an SMS text message with their mobile phone to request the recipient to activate Anne4Care in order to receive a video call. With this extra step, most participants decided to simply use their mobile device instead of Anne4Care to make a video call. In addition, the newspaper and radio functions do not allow participants to search for radio stations, other than the ones preprogrammed.

I cannot find that newspaper, also the radio channel is absent. When I try to search, I get the message ‘no stations available’, so there is nothing programmed I think. [Participant B]

Another challenge was real interactions with Anne4Care. The participants expected the possibility to have a conversation, but that was not possible. Furthermore, part of this interaction was the commands to which Anne4Care often does not react. For example, when a participant asked Anne4Care for the time, or to call someone, Anne4Care may not respond. This could have been caused by the fact that Anne4Care did not recognize all the verbal commands in the Turkish language.

We cannot talk together, because she does not respond. Every morning I hear ‘good morning’, but that is it. [Participant B]

The Anne4Care device itself presented some challenges. One challenge was the anxiety among participants for the devices to overheat, which caused the participants to turn off Anne4Care. Another challenge was the requirement for an internet connection at home. Some of the participants did not have Internet access at the start of the study. Internet access was provided to participants who did not have an internet connection at home during the study period. However, Internet access is not free. Some of the participants are strapped for cash. This makes it difficult for them to pay for Anne4Care and an internet connection when the study ends. They have to make difficult trade-offs.

Now we are using Anne4Care for free, but in a couple of weeks there are probably some costs involved. We do not have Internet connection, I only have a mobile phone subscription with which I am happy. But I am also happy with Anne4Care, so the costs make it quite difficult to make a trade-off. [Participant N]

Because Anne4Care is now only available with an internet connection at home, another challenge is to receive messages from Anne4Care when someone is outside. It would be great, for example, to transfer this with the Anne4Care message app on their mobile phone in case they are not at home. According to some participants, Anne4Care is currently a device only for people who are at home most of the day.

Expectations

This theme represents the ideas, wishes, and future plans of the participants for Anne4Care. In the beginning, some of the participants expressed anxiety toward Anne. They turned the tablet off at night because they thought someone could see or listen to them through the device. One of the new plans most of the participants came up with during the use of Anne4Care was the addition of an option to connect quickly to emergency services.
care services. This connection could be activated by the user, but it should also be activated automatically when older adult immigrants living with dementia do not respond to a call within a period of time. Although all participants expected that Anne4Care would improve the health care of people with deteriorating health or dementia, their expectations were higher at the start.

Anne4Care needs to be improved. At the moment, it is too basic and does not meet the needs of some people. We can do more by ourselves, it is a bit of a disappointment. [Participant D]

As mentioned, it is a challenge to use Anne4Care outside the home because the platform requires the internet. Outside the home, it would be helpful if people could receive medication reminders, therefore, this platform should also be compatible with their mobile devices.

It would be great if I could just take Anne4Care outside. Then I have my medication reminders when I am outside, she will tell me to take the medication and I could take them at the right moment. [Participant M]

There were different ideas for new functions on the current Anne4Care tablet. For example, the addition of short movies or documentaries would allow it to be used for entertainment. Additions to the game function and more options, such as multiplayer games allow an older adult to play with a partner. A range of memory or language-related games, and more challenging puzzles would help meet user preferences. The memory and language games were specifically mentioned by several participants and health care professionals, because of the perceived benefits of cognitive engagement for older adult immigrants living with dementia.

Discussion

Summary of Findings

This study aimed to understand the acceptance of Anne4Care as perceived by older adult immigrants living with dementia using a citizen science approach. This study showed that although older adult immigrants living with dementia had no previous experiences with digital assistive technology specifically, the acceptance of the digital personal assistant, called Anne4Care, by these participants was rather high. Anne4Care was mainly used as a memory assistive tool, for appointments and medication. The use of Anne4Care at first increased the burden of health care professionals and families because health care professionals and sometimes families were responsible for adding new functions as well as appointments in the agenda and changes in the medication list. After the initial effort, most health care professionals and families experienced that Anne4Care reduced their tasks and stress. Contributions of Anne4Care included companionship, help with daily tasks, and opportunities to communicate in multiple languages. On the other hand, some participants expressed anxiety toward the use of Anne4Care and experienced challenges in the use of functionalities, for example, video calling and having real interaction with the avatar. Furthermore, the platform required an internet connection at home and Anne4Care could not be used outside the home.

Reflection on the Literature

The personal digital assistant, Anne4Care, offered companionship for older adult immigrants living with dementia and helped them perform daily activities. Participants mentioned that Anne4Care was very useful as a remember assistive tool, especially for appointments and medications. This finding is consistent with the results of previous studies about technology acceptance in rehabilitation and assistive technologies, and in health care technologies in general [28-30]. Technology acceptance models, for example, the Unified Theory of Acceptance and Use of Technology suggest that if performance expectancy is high, that is, people believe that technologies can help them to achieve their therapeutic goals or achieve their health expectations, this increases the acceptability and actual use of health technologies [31].

A key finding in this study was that participants believed that they did not have all the conditions to use and adopt Anne4Care for a longer period. These conditions are also known as facilitating conditions in technology acceptance theories, for example, the Unified Theory of Acceptance and Use of Technology [31]. Facilitating conditions include, for example, internet connection, technical infrastructure, as well as other internal support such as health care professional involvement and supporting staff (eg, availability of engineers to support the system) [32,33]. In our study, participants reported a lack of internet connection as a main limitation for the acceptance of Anne4Care. There is an extensive body of literature in the field of technology acceptance and use that points toward facilitating conditions as an important determinant factor in technology use [34]. Our result shows that to take full advantage of the potential of digital technologies like Anne4Care, these digital technologies should be accessible also to people of lower socioeconomic status which may have an influence on a person’s autonomy and independence [35].

In our study, health care professionals were involved in using Anne4Care. These health care professionals had close contact and a trustful and respectful relationship with the participants. They played an important role in giving participants information and instruction regarding Anne4Care, in which they can be supportive of the acceptance of technology. Some participants were concerned about their privacy when using Anne4Care. They turned the tablet off at night because they thought someone could see or listen to them through the device. The important role of the health care professionals in this study warrants discussion. The health care professionals, all with a Turkish background, had the expertise to provide access and involve the participants. They were an important frame of reference and guided the participants whenever they received questions, or observed discomfort or doubts. Furthermore, they were key users of health care technology and had a positive attitude toward implementing the new technology. Research suggests that creating a positive, supportive atmosphere is instrumental to the sustainability of technology use [36]. Other studies describe that caregiver engagement is important for the everyday use of technology among people with dementia [33,37].
Although in recent years several technologies have been developed for people with dementia and their caregivers to support self-management, in our study none of the participants knew or used an assistive technology comparable to Anne4Care. All were unfamiliar with the existence of these types of technologies. Globally, there is an increasing aging population and more people staying longer in their own homes which has an impact on society and health care [38]. Although evidence shows that technology for people with dementia seems to stimulate cognitive function and communication skills and reduce loneliness [14,15,39], the implementation of technology can also contribute to the burden of health care professionals and informal caregivers. In our study, participants needed the assistance of health care professionals or families, which increased the demand for health care professionals and families. However, after this initial effort, most health care professionals and families experienced that the tablet facilitated their caregiving roles.

Immigrant populations typically have limited involvement in the development, refinement, and implementation of health care technology. This may be reflected in lower levels of adoption of health care technology. The citizens’ science approach within this study was important for this specific target group. A citizen science approach calls for the optimal involvement of the target group as co-researchers. This extra time and dedication was positively experienced by the older adult immigrant group. It was mentioned that it felt like having a real purpose and gave a sense of fulfilment. This collaboration resulted in an alignment of the technology with the needs and practices of the participants. In future studies, the citizen science approach could also be applied in the development of new technologies as this study focused on the refinement of a technology for a specific target population. Citizen science for health and well-being could provide an effective way to involve vulnerable groups within society to participate in research.

Recommendations for Future Research
Future research could examine ways to increase the implementation of technology among older adult immigrants living with dementia and how to develop the technology competencies of clients, caregivers, and health care professionals. Furthermore, for future research conducting mixed method research using both qualitative and quantitative research methods is recommended to provide more insight into the added value of these kinds of assistive technologies for end users and professionals or family members.

Strengths and Limitations
Although we included only 13 older adults, they were involved in the entire process and collaborated with the researchers as well as with the other participants and their health care professionals. All older adults of the 2 organizations with whom the company Anne4Care started collaboration in the Twente region were invited for this study. The 13 participants who were interested in collaborating were all involved in this study. Another strength of our study was the quadruple collaboration. In our citizen science approach, there was active collaboration between the researchers, older adults, health care professionals, and the company that developed Anne4Care. However, we only involved older adults with an early-stage form of dementia, and 5 older adults were not able to join the second interview, due to COVID-19 illness. In addition, assessing exact levels of health literacy would have provided additional information about the influence of these kinds of assistive technologies on health literacy levels. Finally, a care professional was present and served as an interpreter. This may have impacted the quality of the data collected and nuances in the conversations may have been missed. On the other hand, the presence of the care professional ensured a safe and trustworthy environment.

In this study, a qualitative descriptive research design with a naturalistic inquiry has been used. Four themes were found to be related to acceptance of Anne4Care: (1) use of Anne4Care, (2) positive aspects of Anne4Care, (3) challenges with Anne4Care, and (4) expectations. Data saturation regarding the acceptance of Anne4Care was reached. During the data analysis, similar and confirmation of all findings appeared when analyzing and coding the transcripts. However, besides these 4 themes, there might be other factors relevant to the acceptance of digital personal assistant technology for older persons to stay safe in their homes and be able to age in place.

Conclusions
Although older adult immigrants living with dementia had no experience with digital assistive technology specifically, the acceptance of the digital personal assistant, called Anne4Care, by older adult immigrants living with dementia was rather high. In our study, older adult immigrants living with dementia learned and used Anne4Care with the help of family caregivers. Most older adults accepted Anne4Care into their lives in which Anne4Care offered companionship and helped them to perform daily activities. Older adults provided suggestions for the continued development of Anne4Care.

Acknowledgments
The authors thank the company Anne4Care, which made the device available for the research, and the health care professionals and participants for collaborating with us in every research phase and giving valuable insight. This study was powered by the Twente Regional Deal and received financial support from the Central Government’s Regional Budget, the Province of Overijssel, the Region of Twente, and the Twente Board. The funders had no role in study design, data collection and analysis, the decision to publish, or the preparation of the manuscript.
Authors' Contributions

MB, CMvL, TJJO, ES, and ZM conducted the interviews. MB, CMvL, and TJJO performed the analysis of 2 transcripts and compared codes. Peer debriefing took place at weekly meetings with the project team when scientific and organizational aspects were discussed. All authors contributed to writing the paper and have approved the latest version of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide for first and second interviews with older adults with an immigration background.
[DOCX File, 17 KB - aging_v7i1e50219_app1.docx]

Multimedia Appendix 2
Interview guide for interviews with health care professionals.
[DOCX File, 17 KB - aging_v7i1e50219_app2.docx]

Multimedia Appendix 3
Coding tree.
[DOCX File, 134 KB - aging_v7i1e50219_app3.docx]

Multimedia Appendix 4
Visual information about Anne4Care.
[DOCX File, 727 KB - aging_v7i1e50219_app4.docx]

References


30. Miguel Cruz A, Lopez Portillo HP, Daum C, Rutledge E, King S, Liu L. Mobile app technology acceptance and usability of a mobile app to support the workflow of health care aides who provide services to older adults: pilot mixed methods study. JMIR Aging 2022 May 18;5(2):e37521 [FREE Full text] [doi: 10.2196/37521] [Medline: 35583930]


Edited by M O’Connell; submitted 23.06.23; peer-reviewed by S Snipes, G Goodall; comments to author 21.11.23; revised version received 15.01.24; accepted 21.02.24; published 19.04.24.

Please cite as:
Bults M, van Leersum CM, Olthuis TJJ, Siebrand E, Malik Z, Liu L, Miguel-Cruz A, Jukema JS, den Ouden MEM
Acceptance of a Digital Assistant (Anne4Care) for Older Adult Immigrants Living With Dementia: Qualitative Descriptive Study
JMIR Aging 2024;7:e50219
URL: https://aging.jmir.org/2024/1/e50219
doi: 10.2196/50219
PMID: 38639994

©Marloes Bults, Catharina Margaretha van Leersum, Theodorus Johannes Josef Olthuis, Egbert Siebrand, Zohrah Malik, Lili Liu, Antonio Miguel-Cruz, Jan Seerp Jukema, Marjolein Elisabeth Maria den Ouden. Originally published in JMIR Aging (https://aging.jmir.org), 19.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Technology Usability for People Living With Dementia: Concept Analysis

Shao-Yun Chien1, BSN, MSN, PhD; Oleg Zaslavsky1, BSN, MHA, PhD; Clara Berridge2, MSW, PhD

1School of Nursing, University of Washington, Seattle, WA, United States
2School of Social Work, University of Washington, Seattle, WA, United States

Corresponding Author:
Oleg Zaslavsky, BSN, MHA, PhD
School of Nursing, University of Washington
1959 NE Pacific Street
Seattle, WA, 98195
United States
Phone: 1 2068493301
Email: ozasl@uw.edu

Abstract

Background: Usability is a key indicator of the quality of technology products. In tandem with technological advancements, potential use by individuals with dementia is increasing. However, defining the usability of technology for individuals with dementia remains an ongoing challenge. The diverse and progressive nature of dementia adds complexity to the creation of universal usability criteria, highlighting the need for focused deliberations. Technological interventions offer potential benefits for people living with dementia and caregivers. Amid COVID-19, technology’s role in health care access is growing, especially among older adults. Enabling the diverse population of people living with dementia to enjoy the benefits of technologies requires particular attention to their needs, desires, capabilities, and vulnerabilities to potential harm from technologies. Successful technological interventions for dementia require meticulous consideration of technology usability.

Objective: This concept analysis aims to examine the usability of technology in the context of individuals living with dementia to establish a clear definition for usability within this specific demographic.

Methods: The framework by Walker and Avant was used to guide this concept analysis. We conducted a literature review spanning 1984 to 2024, exploring technology usability for people with dementia through the PubMed, Web of Science, and Google Scholar databases using the keywords “technology usability” and “dementia.” We also incorporated clinical definitions and integrated interview data from 29 dyads comprising individuals with mild Alzheimer dementia and their respective care partners, resulting in a total of 58 older adults. This approach aimed to offer a more comprehensive portrayal of the usability needs of individuals living with dementia, emphasizing practical application.

Results: The evidence from the literature review unveiled that usability encompasses attributes such as acceptable learnability, efficiency, and satisfaction. The clinical perspective on dementia stages, subtypes, and symptoms underscores the importance of tailored technology usability assessment. Feedback from 29 dyads also emphasized the value of simplicity, clear navigation, age-sensitive design, personalized features, and audio support. Thus, design should prioritize personalized assistance for individuals living with dementia, moving away from standardized technological approaches. Synthesized from various sources, the defined usability attributes for individuals living with dementia not only encompass the general usability properties of effectiveness, efficiency, and satisfaction but also include other key factors: adaptability, personalization, intuitiveness, and simplicity, to ensure that technology is supportive and yields tangible benefits for this demographic.

Conclusions: Usability is crucial for people living with dementia when designing technological interventions. It necessitates an understanding of user characteristics, dementia stages, symptoms, needs, and tasks, as well as consideration of varied physical requirements, potential sensory loss, and age-related changes. Disease progression requires adapting to evolving symptoms. Recommendations include versatile, multifunctional technology designs; accommodating diverse needs; and adjusting software functionalities for personalization. Product feature classification can be flexible based on user conditions.

(JMIR Aging 2024;7:e51987) doi:10.2196/51987
KEYWORDS
usability; dementia; older adults; technology; concept analysis; mobile phone

Introduction

Improving Technology Usability for Dementia

Usability, a critical determinant of technology’s quality, influences user acceptance and overall experience, with its importance magnified for individuals with mild cognitive impairment or dementia [1-3]. The intricate nature of modern technologies necessitates designs that accommodate cognitive limitations, ensuring accessibility and ease of use for this demographic. However, the specific usability needs of those with dementia are often not adequately defined given the condition’s variability and progression [4]. This underdefinition underscores the necessity for technology products to be designed with a deep understanding of the cognitive, memory, and learning challenges faced by individuals with dementia. It is critical to ensure that technology is designed not only to be convenient but also to address specific needs and enhance usability across the spectrum of dementia types, thereby facilitating its adoption and effective use.

Worldwide, >50 million people are living with dementia, a number projected to triple by 2050 [5,6]. Every 3 seconds, someone is diagnosed with dementia, posing significant challenges to health care, care provision, and social services worldwide [7]. Technological interventions, including digital tools such as calendars and talking watches, assistive devices, and telecare, offer potential benefits by aiding daily function, improving safety, and enhancing social connectedness, thereby improving life quality and mood [8-11]. Moreover, these interventions can alleviate caregiver burden by providing monitoring and reminders, with a potential to supplement human care [8,12]. The COVID-19 pandemic has further highlighted the crucial role of technology in health care, emphasizing its importance in maintaining quality care [13]. In alignment with the World Health Organization’s “Dementia: a public health priority” report, advancements in communication and assistive technologies have provided a variety of intervention methods to improve the lives and health of people living with dementia, including maintaining independent living and enhancing safety and autonomy [14].

However, technological solutions must be designed with the specific needs and abilities of people living with dementia in mind to avoid confusion, frustration, and potential rejection due to complexity or unintuitive interfaces [15,16]. The World Health Organization’s 2022 policy brief on agism in artificial intelligence (AI) technologies underscores the importance of addressing stereotypes, biases, or discrimination in AI, ensuring quality care for older adults [17]. Tailoring AI technology to the diverse needs, desires, capabilities, and vulnerabilities of people with dementia is essential for harnessing AI benefits while mitigating risks [18]. Thus, considering the usability of technology for dementia care is paramount for successful implementation.

Challenges and Considerations in Technology Usability for Dementia Care

Addressing technology usability for individuals with dementia and their caregivers is recognized worldwide as a priority in national and international funding programs. Despite an increase in information and communications technologies (ICTs) for dementia care, standardized methods for evaluating ICT acceptance and usability for this demographic are lacking [19]. The technology acceptance model (TAM) and its subsequent evolutions, including the TAM 2 and the senior TAM (STAM), highlight perceived ease of use and usefulness as critical to technology adoption [20-23]. These models account for personal factors, such as cognitive abilities, and system features, such as product design and instructional support, emphasizing the adaptability of the latter to enhance usability [24,25]. Therefore, to reduce the digital divide, one of the current promising strategies is to start with product design, adopt user-centered design, and use machine learning techniques to provide timely prompts or suggestions to support and help users [24-26].

The limited adoption of digital technologies by individuals with dementia suggests that usability aspects related to product requirements and design have not been sufficiently addressed [27]. While technology usability research is advancing, there is a notable gap in literature specifically addressing the unique needs of individuals living with dementia [28,29]. This gap in clarity and consistency hampers the advancement of nursing and health science knowledge and the development of theoretical models in technology research [30]. To bridge this knowledge gap and enhance technology usability for dementia care, a concerted effort focusing on clear, tailored design principles and user-centered approaches is crucial.

Aims

This concept analysis sought to examine the usability of technology in the context of individuals living with dementia to establish a clear definition for usability within this specific demographic.

Methods

Overview

Multiple sources were used to delineate the definitions and key attributes of usability specific to people living with dementia. Using the method by Walker and Avant [31], this paper offers a detailed analysis of the essential attributes of usability in the context of dementia. Their methodology uses a 7-step process that facilitates a comprehensive understanding of a concept. The process is structured as follows: (1) select a concept; (2) determine the aim of the analysis; (3) identify all uses of the concept; (4) determine the defining attributes; (5) construct model, borderline, related, and contrary cases; (6) identify antecedents and consequences of the concept; and (7) define the concept’s empirical referents.
For the literature review on technology usability for people living with dementia, works from 1984 to 2024 were searched in the electronic databases PubMed, Web of Science, and Google Scholar. The keywords applied were “technology usability” and “dementia.”

To illustrate the usability considerations of technology for individuals with mild dementia, we drew the interview data from a pilot study conducted by the third author (CB). This study involved the design and evaluation of a web-based interface tailored specifically for this population [32]. This study initially enrolled 33 dyads consisting of individuals living with dementia and their care partners. A total of 4 dyads did not complete the study for various reasons, including health-related issues and the difficulty presented by the standardized survey scales for those with dementia. Consequently, the pilot of the web application was conducted with 29 mild Alzheimer disease dementia dyads, totaling 58 older adults (29 with mild dementia and their respective care partners, with all but 1 pair being spousal dyads). Following the pilot, each dyad was interviewed regarding their experience with the web application. The participant characteristics were as follows. The average age among those living with dementia (n=29) was 70 (SD 7.06) years, ranging from 59 to 82 years, with a male majority (18/29, 62% of individuals). Most were White (28/29, 97%), and 1 participant was African American (3%). The care partners (n=29) had an average age of 68 (SD 6.73) years, ranging from 55 to 83 years, with 38% (n=11) identifying as male. Most were White (27/29, 93%), with 2 (7%) Asian American participants.

Ethical Considerations
The interview data were drawn from a pilot study designed and evaluated by the third author (CB), focusing on a web-based interface specifically for people living with dementia [33]. The study was approved by the University of Washington Division of Human Subjects (STUDY00014226). Informed consent was obtained from all participants. The data are anonymous and have been deidentified. Each individual participant received a Visa gift card for US $150 for their time upon completion of the 3 study components.

Results

Concept Definition
As of now, there has been no comprehensive conceptual analysis concerning the usability of technology for individuals living with dementia. Consequently, to elucidate the use and implications of related concepts, an examination of various sources defining usability, encompassing dictionaries, organizations, and academic studies, is imperative.

Dictionary Definitions
Usability is an abstract and interdisciplinary concept. Merriam-Webster’s Online Dictionary [34] defines usability as “the quality or state of being usable: ease of use.” In the Cambridge Dictionary [35], usability is “the fact of something being easy to use, or the degree to which it is easy to use.” The Oxford English Dictionary [36] offers the following definition: “the degree to which something is able or fit to be used.” These descriptions collectively converge on a comprehensive understanding of usability as a measure of a product’s accessibility and ease of operation. Specifically, when considering individuals living with dementia, the concept of usability is tailored to the degree to which a product is easy to use or suitable for use by people living with dementia.

Definition From the International Organization for Standardization
The introduction section of the guidance on usability by the International Organization for Standardization [27,37] presents one of the most universally acknowledged definitions of usability. According to the International Organization for Standardization standard 9241-11, usability is defined as “the extent to which a system, product, or service can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” (part 11, paragraph 2) [37]. This delineation emphasizes the importance of a user-centered approach in evaluating how well a product or service facilitates the attainment of goals by individuals. Therefore, in assessing usability for people with dementia, the focus is on determining the degree to which the product or service enables effective and efficient goal achievement for this specific user group, underlining the significance of tailoring technology to meet their unique needs and enhance their quality of life.

Literature Definitions
While the studies reviewed did not offer explicit definitions of “usability for individuals living with dementia,” the prevailing focus within the literature revolves around the examination and assessment of technology interventions tailored for this demographic. These interventions encompass a spectrum of tools and methodologies designed to gauge acceptance, adoption, and usability among individuals living with dementia. Thus, by analyzing the content of these articles, we can derive insights into the underlying characteristics and properties associated with usability in this context. The concept of technology usability integrates essential characteristics, including ease of use, satisfaction, learnability, utility, effectiveness, efficiency, flexibility, familiarity, responsiveness, and the clarity and visibility of feedback mechanisms [38]. This comprehensive understanding of usability encompasses both objective and subjective elements, considering objective performance indicators such as actual use efficiency, effectiveness, and error rates alongside the subjective user experience and perceptions [4].

Research indicates that 40% of individuals with dementia require additional assistance and time to understand and use technological tools, including navigation through certain icons and devices. This underscores the necessity for facilitators or supervisors to aid in explaining operational steps and the use
of technological intervention tools [4]. The analysis by Shultz and Hand [30] enriches our understanding of usability as the degree to which technology is perceived by users as learnable, efficient, and satisfactory. In parallel, Gibson et al [41] defined usability as the degree to which the user perceives acceptable learnability, efficiency, and satisfaction when using the technology. Gibson et al [41] advocate for a design philosophy that emphasizes personalized and adaptable support for individuals living with dementia, advocating against the use of generic technological solutions.

Further elaborating on these insights, Asghar et al [42] articulate how assistive technologies enable social interaction, health monitoring, independent mobility, and punctual medication adherence, thereby supporting daily living activities and enhancing the quality of life of those living with dementia. This perspective reiterates usability’s core definition as ease of use, efficiency, and the capacity to meet specific user needs.

The investigative effort by Miguel Cruz et al [19] into the assessment methods for the acceptance, adoption, and usability of ICTs by people living with dementia and their caregivers highlights the critical gap in standardized evaluation methodologies. This gap signals the need for further research and development to measure such technologies effectively. Hence, usability for people living with dementia hinges on the extent to which they can use specific technologies effectively, satisfyingly, and safely considering their unique requirements and cognitive limitations.

In the domain of mobile app development [43], the initiative to enhance the quality of life of those living with dementia and Alzheimer disease by creating older adult–friendly apps is paramount. This initiative directly addresses the cognitive and usability challenges faced by this demographic, aiming to make technology both accessible and beneficial.

The usability of mobile apps is fundamentally tied to 9 thematic areas: user interface design, physical considerations, screen size, interaction challenges, meeting user needs, addressing the lack of self-awareness regarding app necessities, mitigating the stigma associated with app use, overcoming technological inexperience, and emphasizing the importance of technical support [32]. These areas highlight the critical need for developing intuitive, user-friendly apps tailored to the unique challenges encountered by individuals with dementia, thereby significantly enhancing their autonomy and quality of life.

Considering the specific needs and limitations of individuals with dementia or mild cognitive impairments when designing and implementing technological interventions is critical. Through appropriate support and assistance, their user experience and the effectiveness of the intervention can be improved [4]. Usability is related to user friendliness and ease of use and learning, serving as a means for older adults and those with reduced capabilities to participate in activities and engage equally in society. Highlighting the value of involving users in technology development and clinical trials, the design of intervention studies should include people with dementia and their caregivers to understand the design features necessary to enhance usability and acceptance [9].

Kung and Chen [28] conducted a concept analysis on the usability of health promotion mobile apps, summarizing the characteristics of usability. The most common attributes of usability identified in their study included efficiency, user satisfaction, and learnability [28]. These attributes are crucial for health promotion apps, ensuring that user expectations are met, providing satisfaction, and facilitating ease of learning and use. Textbox 1 organizes the characteristics of usability for people living with dementia from the literature review. Notably, ease of use, effectiveness, efficiency, and satisfaction were frequently mentioned across various studies; these attributes are generally recognized as core characteristics of usability across various fields, not just in the context of designing for people living with dementia. Other significant attributes that repeatedly appear in the literature on usability for people living with dementia include adaptability, personalization, intuitiveness, and simplicity, suggesting a comprehensive approach to addressing the unique needs and challenges faced by people living with dementia.
Textbox 1. Summary of the characteristics of the definition of usability for people living with dementia.

<table>
<thead>
<tr>
<th>Sources and attributes of usability for people living with dementia</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Boulay et al [44], 2011: effectiveness, efficiency, and satisfaction</td>
</tr>
<tr>
<td>• Lim et al [45], 2012: intuitive, engagement, and adaptability to users’ needs</td>
</tr>
<tr>
<td>• Meiland et al [39], 2012: user friendliness, usefulness, and effectiveness</td>
</tr>
<tr>
<td>• González-Palau et al [46], 2013: effectiveness, efficiency, and satisfaction</td>
</tr>
<tr>
<td>• Yamagata et al [43], 2013: ease of use, personalization, accessibility, adaptability, and engagement</td>
</tr>
<tr>
<td>• Boger et al [47], 2015: intuitiveness, simplicity, customization, and adaptability</td>
</tr>
<tr>
<td>• Manera et al [48], 2015: effectiveness, efficiency, satisfaction, and engagement</td>
</tr>
<tr>
<td>• Martins et al [38], 2015: ease of use, satisfaction, learnability, efficacy, coherence, flexibility, and responsiveness</td>
</tr>
<tr>
<td>• Lindqvist et al [40], 2015: ease of use, effectiveness, and adaptability to users’ needs</td>
</tr>
<tr>
<td>• Shultz and Hand [30], 2015: learnability, efficiency, and satisfaction</td>
</tr>
<tr>
<td>• Ben-Sadoun et al [49], 2016: effectiveness, efficiency, and satisfaction</td>
</tr>
<tr>
<td>• Garcia-Sanjuan et al [50], 2017: effectiveness, efficiency, satisfaction, and engagement</td>
</tr>
<tr>
<td>• Meiland et al [51], 2017: effectiveness, efficiency, simplicity, intuitiveness, personalization, and engagement</td>
</tr>
<tr>
<td>• Tziraki et al [52], 2017: engagement, acceptability, and accessibility</td>
</tr>
<tr>
<td>• Asghar et al [42], 2018: effectiveness, adaptability, and satisfaction</td>
</tr>
<tr>
<td>• Holthe et al [9], 2018: effectiveness, efficiency, and satisfaction</td>
</tr>
<tr>
<td>• Gibson et al [41], 2019: adaptability, accessibility, and effectiveness</td>
</tr>
<tr>
<td>• Tuen et al [53], 2020: ease of use, intuitiveness, effectiveness, engagement, personalization, and adaptability</td>
</tr>
<tr>
<td>• Contreras-Somoza et al [4], 2021: effectiveness, efficiency, and satisfaction</td>
</tr>
<tr>
<td>• Engelsma et al [54], 2021: personalization, simplicity, clarity, engagement, and adaptability</td>
</tr>
<tr>
<td>• Koh et al [55], 2022: ease of use and engagement, customizability, and adaptability</td>
</tr>
<tr>
<td>• Neubauer et al [56], 2022: effectiveness, efficiency, simplicity, and satisfaction</td>
</tr>
<tr>
<td>• Miguel Cruz et al [19], 2023: simplicity and adaptability</td>
</tr>
<tr>
<td>• Ye et al [32], 2023: ease of use, accessibility, personalization, and effectiveness</td>
</tr>
<tr>
<td>• Zheng et al [57], 2023: effectiveness, efficiency, satisfaction, simplicity and clarity, and engagement</td>
</tr>
<tr>
<td>• Zhu et al [58], 2024: ease of use, accessibility, personalization, and engagement</td>
</tr>
</tbody>
</table>

**Potentially Related but Distinct Concept of Acceptability**

Acceptability of technology is the extent of the primary users’ predisposition to implement the technology in their daily activities as a result of their diverse perceptions of the product’s characteristics [59]. The primary users of technologies for people living with dementia may include individuals living with dementia or their care partners. Some studies have reported the technology acceptability of people with mild to moderate dementia and their caregivers. For example, an inaesthetic device may be interpreted as unacceptable because of its unappealing appearance and, therefore, has lower acceptability.

**Describing Usability for People Living With Dementia From the Clinical Standpoint**

Dementia progresses slowly into 3 stages: mild, moderate, and severe, which can also be called early, middle, and late stages [60]. According to the Clinical Dementia Rating (CDR), there are different stages of dementia; based on an individual’s cognitive functions, a 5-point system is used to describe each stage of dementia. A person in CDR-0 does not have dementia, and stage CDR-0.5 is considered very mild dementia or mild cognitive impairment, with slight but consistent memory problems. The expected duration of this stage is 3 to 7 years. Someone with CDR-1 has mild dementia. The average duration of this stage is 2 years, with memory loss, particularly recent events, and having trouble with the inability to perform daily tasks. In stage CDR-2, a person living with moderate dementia experiences more profound memory loss and is typically disoriented to time and place for the expected duration of 2 to 4 years. A person in stage CDR-3 is considered to have severe dementia. One might have multiple medical comorbidities, which often result in the end of functional independence. The average duration of stage CDR-3 is typically 1 to 2.5 years [60-63]. These stages help us understand how dementia may change over time and serve as a guide for the design of technology products at different stages, addressing different needs and usability.
Interactive technology refers to systems and devices that can respond to user inputs in real time, allowing for a dynamic exchange of information between the user and the technology. This domain of technology is characterized by the bidirectional provision and reception of information through actions such as language, text, movement, and touch. The key features of interactive technology include user input, real-time processing, and outputs that adjust based on user actions. Users communicate their requirements, and the technology delivers the results of their operations back to them through the interface [64, 65]. Common forms of interactive technology, such as touch screens on smartphones, tablets, and information kiosks, as well as virtual reality, augmented reality, smart home devices, social media, and video games, enable users to directly interact with displayed content. Noninteractive technologies refer to products that do not require or allow for instant user input or feedback to operate. Unlike interactive technologies, which are designed to involve users in 2-way communication or interaction, noninteractive technologies, once activated or triggered, operate independently of user inputs. These technologies are often designed to perform specific tasks, display information, or execute commands without requiring ongoing user interaction. Television, smart speakers, sensors, static websites, and radio are typical examples of noninteractive technology. Operating interactive technology typically requires hand-eye coordination and precise control of the product, as well as reading or memorizing the steps or instructions on how to operate the technological product. It may even involve understanding complex user interfaces or executing multistep tasks to achieve the intended use of the product. Compared with the middle and late stages, people with mild dementia are still able to retain independence to complete many daily activities or do so with little assistance or guidance [60, 66], and they are more likely to use interactive technology. For example, people living with mild dementia who attempt to have cognitive training using interactive tools to stimulate their cognitive functions might have one set of usability needs. The course of dementia is progressive, and when it enters the middle stage, people living with dementia have more profound memory loss. Their ability to cope with daily life becomes more complicated [60, 66]. Due to the decline in memory functions and the ability to perform fine movements, the demand for technology products in the mid to late stages may be more often focused on noninteractive products due to usability issues.

In addition, different individuals may have different symptomology. There are different subtypes of dementia. Alzheimer disease and Lewy body dementia are the 2 most common progressive dementias. In addition to cognitive impairment, Lewy body dementia may be accompanied by muscle stiffness, hand tremors, and unsteady gait at an early stage [66]. A person with Lewy body dementia might have mobility issues that might imply the usability concern of motor activity, and someone with Alzheimer dementia might have issues with remembering instructions.

In short, from a clinical point of view, to look at the technology usability for people living with dementia, many factors require assessment, such as the stage, type, and symptoms of dementia. Furthermore, the categories of technology—interactive or noninteractive technology—should be considered based on the characteristics and needs of the disease, focusing precisely on the needs of tasks to be performed (Figure 1). In addition, people living with dementia retain certain relevant functions that are crucial for designers to consider. Despite the cognitive decline associated with dementia, individuals often maintain sensory and motor abilities provided they do not have other chronic conditions or age-related changes [67]. For designers, it is essential to recognize and leverage these retained functions when developing tools for this population. For instance, people living with dementia may be able to recall distant personal memories, which can be used in product design to evoke positive emotional and behavioral outcomes. By understanding and designing with knowledge of these maintained functions, digital tools can be better aligned with the abilities and needs of people living with dementia, thereby enhancing their engagement, enjoyment, and overall quality of life [67].

https://aging.jmir.org/2024/1/e51987

( PAGE NUMBER NOT FOR CITATION PURPOSES)
**Definition From Interviews With People Living With Dementia**

To illustrate technology usability considerations in people with mild dementia, we drew from a pilot study in which the third author designed and evaluated a web-based interface specifically for this population. In brief, the project aimed to investigate the experience of using a web application interface designed for people living with dementia and their care partners [33]. Its design complied with the criteria of the Web Content Accessibility Guidelines 2.1 at level AA, which is the standard for web accessibility developed by the World Wide Web Consortium. Components to enhance accessibility for people with sensory loss and with dementia included an audio option; screen reader compatibility; and clear, short sentences that avoided abstractions and multiple constructs. After conducting a pilot of the web application with 29 dyads of people with mild Alzheimer disease dementia and their care partners for a total of 58 older adults, each dyad was interviewed about their experience using it together. The interviewees indicated what they found positive about the web application that contributed to usability. They stated that they liked that the web application’s interface was simple, clean, straightforward, and easy to use and that there were not too many options. The buttons were specific and clearly visible, and there were not too many distractions on the screen to create confusion. Participants appreciated how age-related vision changes were considered in the design. For example, the buttons on the web application were large enough, and the texts were of a size suitable for reading. Furthermore, because they could forget the operation steps or the functions of unmarked buttons, they found the clear and explicit wayfinding directions helpful and felt that they did not risk “getting lost” in the web application. Pilot participants further appreciated personalization options, such as open-text boxes to type in explanations so that the information could be entered in a familiar way. In addition, users said that they liked the app’s summary feature, which reports the answers they chose for their future reference.

Many dyads of people with dementia and their care partners said that it was helpful to have an audio option, which read the text on the screen when selected while visually highlighting each sentence. It assisted people living with dementia in concentrating, and the care partners enjoyed not needing to read the text on the application aloud to their partners. Approximately
1 in 3 older adults experience age-related hearing loss, and their hearing sensitivity, especially high-frequency hearing, decreases gradually [68,69]. However, the ability to hear low-pitched sounds is often not affected [70]. Therefore, the web application used a male-sounding (lower) voice for reading, which can be heard more clearly by older adults. This and other age-related physiological changes are factors that must be taken into account as dementia is associated with older age [5].

Finally, in the usability-testing phase, the pilot found that it was sometimes challenging to obtain feedback from people living with dementia on the web application within a few days due to short-term memory difficulties. Some participants noted that they were unable to remember details of their experience or the web application well enough to report on it. Hence, to understand the user experience, it is necessary to find creative solutions such as soliciting real-time feedback during or immediately after using the application.

Defining Attributes

Overview

Following the approach by Walker and Avant [31], after the concept of usability is defined, its essential attributes must also be defined. Particular attributes are mentioned repeatedly in the literature and can be used to distinguish this concept from those that are similar or related. The current definition and characteristics of usability are not entirely applicable to individuals with dementia. Therefore, this study synthesized a definition of usability for patients with dementia based on multiple resources, including dictionaries, a literature review, and data from interviews with people living with dementia and their caregivers. In addition to ease of use, effectiveness, acceptable learnability, and satisfaction, special attention is required for dementia symptoms and aging. Personalization becomes a crucial attribute, with adaptations based on dementia stages, subtypes, and symptoms. It emphasizes the value of simplicity, clear navigation, age-sensitive design, and personalized features, making the design of the technology or interface more specific and intuitive.

Constructed Cases

The application of technology to people living with dementia can be divided into the following five categories: (1) cognitive training and daily living, (2) screening, (3) health and safety monitoring, (4) leisure and socialization, and (5) navigation [71]. Selected clinical cases are listed herein to clarify and illustrate the concept of usability for people living with dementia. The case scenarios presented in this study are a combination of adapted cases from the first author’s previous clinical work and fictional scenarios, and all mentioned names are of a fictional nature. On the basis of the work by Walker and Avant [31], the exemplar case contains all the defining attributes of the concept. The borderline case possesses not all but most of the defining attributes of the concept, and the related case is related to the main defining attributes but does not contain all of them. In addition, a contrary case that has no defining attributes of this concept but is contrary to it is presented.

Exemplary Case

Mr Wang is a patient who was diagnosed with mild dementia 6 months ago. He experiences memory impairment, has difficulties in concentration, becomes lost easily, frequently forgets to lock the door and turn off the tap, is unable to plan trips or prepare meals as he once could, and often feels depressed. Nevertheless, he can maintain his daily life. To delay deterioration, stimulate the brain, and relieve emotional distress, his nurse provided him with a tablet computer. A total of 11 apps were preinstalled on the tablet computer based on 3 categories. The creative (art and music) app enables users to produce artwork on a blank canvas by accessing various brush types and a wide array of colors. Simple interactive games allow users to interact with virtual animals that respond to a user’s touch and gestures. Relaxation apps allow users to play relaxing background music while the tablet screen displays nature pictures [45].

After 30 minutes of demonstration, guidance, and practical exercises by the nurse, Mr Wang learned how to operate the device, including turning it on and off, charging it, adjusting the volume, navigating the home screen, and starting and exiting apps. He also received a cell phone to contact someone in case of any operation problems. After completing the standard training checklist, Mr Wang returned home with the tablet and used it for 7 days to evaluate it and the apps’ effectiveness. The nurse invited Mr Wang’s primary caregivers to observe the status of his use of the device. Mr Wang used the tablet computer for an average of 50 minutes per day, which exceeded the recommended 30 minutes. Mr Wang expressed satisfaction with using the tablet and the apps and was willing to continue using it. According to his caregivers’ observations, Mr Wang spent more time on interactive gaming and listening to music. He spent an average of 15 minutes independently using the tablet, and he was able to independently store and charge it. Use of the tablet also enhanced his emotional stability.

Borderline Case

A person living with mild dementia tried to use a mobile app to buy train tickets, and their caregiver, also an older adult, assisted him in its operation (co-use). However, the app’s user interface was complicated to operate. They were finally able to buy the ticket using the app, but the process was too time-intensive. This scenario involves a user living with dementia, technology, tasks to be performed (buying train tickets), and the consequences of co-use with the patient’s caregiver. The patient used the app several times afterward and became more proficient, indicating that the operation of the app is learnable for a person living with dementia. Moreover, compared with visiting the train station to purchase train tickets directly, using the app to buy tickets is more efficient. Although the purpose of using this tool was achieved, that is, it is applicable, the patient was dissatisfied with the complex and time-intensive user experience; therefore, the satisfaction attribute was not met.

Related Case

Mrs Smith has a diagnosis of mild dementia and lives with her family. Her daughter assisted her in downloading a diet and
fitness app on her tablet computer to track nutrition, fitness, and health data. The app offered a menu and colorful photos of various diet types as well as daily exercise videos to follow. In addition, users could upload pictures and share their progress and weight control results in the app’s discussion forum, where other users could “like” and comment on the posts.

Mrs Smith enjoyed using the app for its vivid food pictures, animated exercise videos, and users sharing positive posts related to fitness progress on the forum; she felt delighted and vibrant when “liking” forum posts. For Mrs Smith, this app was useful, engaging, and a source of pleasure in her life. Hence, usefulness was present as a defining attribute. However, she could not comprehend how to record diet and health information because the operation interface was too difficult to use; the app lacked the learnability attribute. In addition, the app did not achieve the expected outcomes of its design (healthy diet and fitness), although effectiveness was not included as a defining attribute in this case. Despite this, Mrs Smith was satisfied with using this app, fulfilling the satisfaction attribute.

Contrary Case
A patient with severe dementia lives in a long-term care facility and is cared for by nursing staff. Because of his age and symptoms, this patient has a high risk of falling. The staff assisted in outfitting the patient with a medical alert necklace that automatically sent messages to set recipients if the patient had an emergency, such as a fall. Because this product has an automatic detection system, the patient is not required to learn how to use it, and no learnability attribute is present. The patient himself was unable to understand the function and benefits of the product as a result of the deterioration of his cognitive function; he thought something uncomfortable was stuck on his neck and often removed the necklacelike detector. Despite staff perception that it was essential to his personal safety, he was dissatisfied with it. He often removed the detector and failed to carry it, reducing the usefulness and effectiveness of the product.

Antecedents
The antecedents of usability for people living with dementia encompass a complex interplay among the individual living with dementia, the technology in use, the tasks to be performed, and the concept of co-use. The usability concept for people living with dementia is profoundly influenced by a complex interplay of social, environmental, and personal factors. The unique needs and capabilities of people living with dementia present various variables that affect usability, making the user living with dementia the most relevant antecedent. However, beyond the individual user, it is essential to consider the broader context of technology use. This includes the availability of space, the skills and knowledge required to operate the technology, existing support systems (such as caregivers or health care professionals), and the financial resources needed to access the technology. Furthermore, depending on the severity of dementia, patients may face greater challenges in using technological tools independently, necessitating co-use or shared use with their caregivers [72]. The concept of “co-use” underscores the importance of social dynamics in usability.

Consequences
The consequences of people living with dementia using technology include co-use, usefulness, and satisfaction. As progressive and irreversible cognitive impairment can affect the capability of people living with dementia to operate technological products, being able to co-use the products is one of the criteria of the consequences of the concept of technology usability. On the basis of the varying degrees of dementia severity, patients who exhibit more pronounced symptoms often encounter greater challenges in independently using technological tools. This situation necessitates a collaborative approach where co-use or shared use with caregivers becomes essential. The need for such support intensifies as the severity of the dementia increases, highlighting the importance of adaptable technological solutions and the involvement of caregivers in the use process to ensure effective and safe interaction with these tools [72]. In addition, the research by Astell et al [73] on the use of interactive multimedia touch screen systems for people living with dementia found that interacting with caregivers through touch screen tablets was intriguing and engaging for people living with dementia. These consequences are similar to the attributes that define usability. Nevertheless, the consequence of usability is the result of the direct evaluation of specific technological applications by people living with dementia, including co-use, perceptions of usefulness, and satisfaction [30] (Figure 2).

Figure 2. Overview of antecedents, attributes, and consequences of technology usability for people living with dementia.
Empirical Referents

Overview

A deeper understanding of the concept of technology usability for people living with dementia can be used as a reference index for clinical practice in assisting those living with dementia in locating suitable technological interventions that promote a healthy and comfortable life. Although the exact definitions and testing of usability vary for different products, their effectiveness is reflected in the actual user operation, that is, whether the device achieves its main goal and is satisfactory for the users (people living with dementia). When testing usability, certain elements must be considered, such as the target users; their learning abilities, desires, and needs; the environment in which they operate; and what advantages the product offers [73]. Usability evaluation methods include quantitative methods such as questionnaires and task completion and qualitative methods such as the “think-aloud” protocol, focus groups, and interviews [74].

Think-Aloud Methods

Among the evaluation methods, the think-aloud (or thinking-aloud) protocol is the most commonly applied qualitative method used to collect data in usability testing. Users are required to speak their thoughts, actions, and perceptions aloud while operating a product, with observers objectively recording the user’s speech verbatim without interpreting their behaviors or statements [75]. Customarily, the test is filmed for developer review; developers observe the user process of operating the product as well as users’ response to ascertain their experience of the application [76].

System Usability Scale

This is the most widely used and validated questionnaire and comprises a 10-item Likert scale. This simple, standardized questionnaire is advantageous for its objectivity, generality, repeatability, and quantifiability. The questions focus on frequency of use, complexity of operation, consistency, degree of preference, and achievability of the product goal [77]. However, because the think-aloud protocol has higher demands of cognitive load for people living with dementia [76] and the System Usability Scale, a posttest tool, cannot reflect how people living with dementia interact with technology products timely, there is currently no clear, standardized method to assess usability for people living with dementia. Hence, there is still a higher need to design more consistent and reliable evaluation instruments for people living with dementia [19,76].

Observation and Logging

Other primary methods for evaluating usability for people living with dementia are observational methods and logging, which present more objective data, with observations enhanced by video recordings for more reliable results [32,76]. However, observer bias, the presence and perceptions of the observer, can influence both the observer’s interpretations of what is seen and possibly the behavior of the participants being observed.

Therefore, it is advised to combine both objective and subjective measures for a comprehensive understanding of usability needs [78,79] and consider the aims of the research to select the proper measuring tool.

We suggest that, when considering usability for people living with dementia, it is necessary to understand the characteristics of users. First, the symptoms and needs of dementia at each stage, the designing purpose, and the tasks to be performed should be understood. Furthermore, dementia mostly affects older adults, so corresponding designs should take into account diverse physical needs, potential sensory loss, and other age-related changes. In addition, primary caregivers spend a lot of time and energy accompanying and caring for people living with dementia. Their firsthand observations, insights, and opinions can be informative when testing products, particularly for technology products designed for co-use. As the disease progresses, changes in condition and symptoms need to be considered. Another suggestion is to design diversified and multifunctional technological products to provide more inclusive and flexible choices and adjust software features to meet individual needs. Product function classification could be upgradable or downgradable as the condition of the user changes, and these functions could be selected by the user.

Furthermore, Sebastian et al [80] have raised a question about the potential of rhetoric in enhancing the adoption of AI; their study results suggest that the adoption of strategic communication techniques (ethos, pathos, and logos) can significantly impact people’s willingness to accept AI technologies. Our conceptual analysis also acknowledges the importance of effective communication strategies in facilitating technology adoption. This insight is particularly significant in technology adoption among individuals living with dementia, where trust in and understanding of technology plays a crucial role. Therefore, future research could explore how customized communication strategies based on rhetorical principles can address the concerns and needs of people living with dementia and their caregivers.

Discussion

Clarifying and Validating the Concept of Usability for Individuals With Dementia

To our knowledge, this study represents the first concept analysis specifically focused on usability for individuals with dementia. This paper integrates existing literature and combines empirical data obtained from interviews to relate the concept to specific real-world situations. It offers insights into usability within the context of dementia, covering its practical significance and applications as well as directions for future research.

The concept of usability, particularly for individuals living with dementia, demands a nuanced understanding that accounts for the rapid advancements in technology and its increasing application in supporting these individuals and their caregivers. The specific needs of individuals with dementia, tailored to accommodate common symptoms and optimize their capabilities, underline the critical need for a precise and comprehensive definition of technology usability within this context.
Usability is acknowledged as a multidimensional construct that becomes tangible only when technology is actively used by individuals. The variability of the usability definition, contingent on the context and specific application field, presents a challenge in achieving a unified conceptualization [19,81]. This challenge is further compounded for individuals living with dementia due to the complex interplay between the severity and subtype of dementia and specific technological requirements, including the necessary level of interactivity, technical characteristics, and intended tasks.

Given the diverse severities and progression rates of dementia, individuals experience varying needs across different stages of the condition, necessitating distinct technological interventions and, consequently, different usability requirements at each phase. Therefore, as delineated in Textbox 1, beyond the standard usability attributes—ease of use, effectiveness, efficiency, and satisfaction—additional attributes pertinent to individuals living with dementia include intuitiveness, simplicity, personalization, and adaptability.

Intuitiveness refers to the ease with which users can understand and interact with technology or an interface without previous instruction. When technology is intuitive, individuals living with dementia are more likely to use it independently, fostering a sense of autonomy and confidence. Simplicity, on the other hand, emphasizes clean design and straightforward functionalities that avoid overwhelming the user. It ensures that the cognitive load is minimized, which is particularly important given the cognitive challenges associated with dementia. Together, these attributes create a user-friendly environment that supports the engagement and sustained use of technology by minimizing frustration and maximizing ease of use.

Personalization significantly enhances usability and the overall user experience. Although not a traditional usability attribute, it reflects an understanding of the variability in dementia symptoms and stages, requiring solutions tailored to the individual’s changing needs. This approach specifically addresses the challenges associated with various stages and types of dementia. Adaptability complements this concept by ensuring that technology can adjust to the user’s changing condition over time, particularly as dementia progresses.

**Exploring the Integration With Existing Theoretical Frameworks**

The TAM posits that perceived ease of use and perceived usefulness are the 2 main factors affecting one’s beliefs, intentions, and attitudes toward using novel technology [20,21]. These factors become even more significant in the context of technology development for patients with dementia, where cognitive impairments necessitate a more detailed and specific consideration of usability. This underscores the need to optimize technology products based on the specific abilities of users. Venkatesh and Davis [22] developed the TAM 2, an expansion of the original TAM, which highlights the importance of social influence and cognitive instrumental processes in technology acceptance and use, including aspects such as subjective norms, voluntariness, and individual experiences. These factors may play a crucial role in determining whether patients with dementia are willing to adopt certain technologies. However, for individuals with dementia, the usability of technology products is particularly critical as they may encounter specific challenges in cognition and perception that are not as pronounced in the general user groups highlighted by the TAM and TAM 2. This means that, for this unique population, perceived ease of use and perceived usefulness remain critical factors but there is also a need to further consider how products can be designed and supported to meet their specific needs.

The STAM, built on the foundation of the TAM and specifically designed to address the needs of older adults, including those with dementia, focuses on technology adoption among older adults [23]. The STAM maintains that perceived usefulness and ease of use, key factors in technology adoption across different age groups, are especially critical for older adults. Moreover, it adds that gerontechnology self-efficacy, gerontechnology anxiety, and facilitating conditions, which are age-related health and ability characteristics more relevant to older adults, are vital in designing technology products for patients with dementia. These considerations emphasize the necessity of considering their cognitive limitations and technology use experience as well as how to minimize barriers to technology use through facilitating conditions.

Therefore, based on the usability for people living with dementia, it is suggested to develop or expand the existing TAMs and theories by (1) incorporating specific factors related to cognitive impairments, adding variables related to cognitive load, information processing speed, and memory retention capabilities; (2) considering contextual and environmental factors such as the home environment, support from care institutions, and the impact of social and cultural backgrounds on technology acceptance and use; (3) emphasizing the importance of user experience design, especially in terms of how design can reduce anxiety and boost confidence among users with dementia, thereby promoting technology acceptance [4]; and (4) focusing on personalized and adaptive design for patients with dementia, expanding the model to include principles of adaptability and personalization according to individual differences, considering the varying needs of people at different stages of dementia [32].

**Recommendations for Standardized Usability Testing for People Living With Dementia**

We recommend that, when developing a standardized usability test for people with dementia, it is essential to integrate several elements, particularly considering the progression and variability of symptoms. In addressing the usability attributes relevant to people living with dementia, the use of adaptive testing methods is advised. These methods involve designing usability tests that can be customized for different stages of dementia, adjusting protocols based on the cognitive, physical, and emotional states of participants. Due to the progressive nature of dementia, regular re-evaluation is necessary. Immediate feedback after testing is important due to short-term memory difficulties in people living with dementia. To prevent attention deficits or fatigue in people living with mild to moderate stages of dementia, testing durations should be limited as participants might become distracted and fatigued after 30 to 35 minutes of continuous dialogue [33]. Inclusive design principles must...
consider a broad range of abilities and limitations, ensuring test accessibility for various dementia stages. Finally, collaboration with experts such as health care professionals and dementia specialists is vital for deeper insights into the needs and challenges of people living with dementia.

### Potential Frameworks for Usability Assessment for People Living With Dementia

Furthermore, given the unique needs and challenges faced by people living with dementia, addressing the need for a standardized usability assessment specifically tailored to them is important. Some potential frameworks or methods that could be considered or developed for this purpose include, first, human-centered design approach. In this approach, the focus is on involving people living with dementia and their caregivers in the design process by adopting a human-centered design for the intervention, involving an interactive development process that focuses on the users and their needs and requirements [82]. This could include interviews, focus groups, and usability testing sessions with prototypes [74]. Participants are the ones who understand their needs best; using this method, they will also be the designers, involved in the designing phase. It is crucial to incorporate their feedback into the development process to tailor the product to meet their specific capabilities, needs, and preferences. The second potential method is contextual design and observational studies. This method consists of observing people living with dementia interacting with technology in their usual environment. It helps understand how they use technology, what challenges they face, and what aspects of the technology are most beneficial or problematic for them [83]. The third method is heuristic evaluation, a method in which usability experts analyze a product using a set of 10 heuristics [84] and that can be adapted for people living with dementia by involving dementia specialists and usability professionals [85]. The fourth method is longitudinal studies. As dementia is a progressive condition, longitudinal studies can be valuable in assessing how usability needs change over time and how well technology adapts to these changing needs.

### Limitations and Future Research Directions

Despite obtaining rich information through interview data, this research still has limitations. One of the limitations is the high homogeneity in terms of the participants’ race and ethnicity, which could potentially impact the generalizability of the research findings. In addition, it is suggested that future research collect data on participants’ educational background and socioeconomic status as these factors may be related to their willingness and ability to use technology products.

### Conclusions

This study, through a concept analysis that included interviews with people living with dementia and their caregivers, clarified the definition and attributes of usability tailored for this population. A precise definition of usability in this context is crucial to guide future research and practical applications. This study stresses the importance of considering dementia-specific aspects such as the symptoms and aging process. Customization, guided by the disease’s stages, subtypes, and symptoms, is emphasized as critical. Therefore, the design of interventions for individuals with dementia should prioritize simplicity, clear navigation, age-appropriate esthetics, and personalization to enhance specificity and intuitiveness. Furthermore, it is necessary to consider the antecedents, attributes, and consequences of technology usability for this demographic. Adopting a comprehensive approach is pivotal for developing technology solutions that are finely attuned to the unique needs of people living with dementia, fostering a nuanced understanding of usability in this context.

### Acknowledgments

The authors wish to express their sincere gratitude to the McLaw Scholarship, the University of Washington School of Nursing Digital Health Innovation Hub, and the Aljoya Endowed Professorship in Aging for their generous support. The dyadic pilot study was supported by the Alzheimer’s Disease Research Center of the University of Washington [P30AG066509] and the National Institute on Aging (K01AG062681). The financial assistance provided by these entities has been crucial in advancing the authors’ research and enriching their contributions to the field. They also extend their appreciation to the reviewers, whose insightful comments and suggestions significantly enhanced the quality of this manuscript.

### Data Availability

The data sets analyzed during this study are not publicly available due to the reuse restrictions outlined in the consent form.

### Conflicts of Interest

None declared.

### References


62. Dementia Care Central. URL: https://www.dementiacarecentral.com/aboutdementia/facts/stages/ [accessed 2023-12-23]


Abbreviations

AI: artificial intelligence
CDR: Clinical Dementia Rating
ICT: information and communications technology
STAM: senior technology acceptance model
TAM: technology acceptance model

©Shao-Yun Chien, Oleg Zaslavsky, Clara Berridge. Originally published in JMIR Aging (https://aging.jmir.org), 03.07.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Review

Characteristics and Range of Reviews About Technologies for Aging in Place: Scoping Review of Reviews

Jenny M Bergschöld1, PhD; Mari Gunnes1, PhD; Arne H Eide2, PhD; Eva Lassemo1, PhD

1Department of Health, SINTEF Digital, Trondheim, Norway
2Department of Health, SINTEF Digital, Oslo, Norway

Corresponding Author:
Jenny M Bergschöld, PhD
Department of Health
SINTEF Digital
Professor Brochs gate 2
Trondheim
Norway
Phone: 47 46314138
Email: jenny.bergschold@gmail.com

Abstract

Background: It is a contemporary and global challenge that the increasing number of older people requiring care will surpass the available caregivers. Solutions are needed to help older people maintain their health, prevent disability, and delay or avoid dependency on others. Technology can enable older people to age in place while maintaining their dignity and quality of life. Literature reviews on this topic have become important tools for researchers, practitioners, policy makers, and decision makers who need to navigate and access the extensive available evidence. Due to the large number and diversity of existing reviews, there is a need for a review of reviews that provides an overview of the range and characteristics of the evidence on technology for aging in place.

Objective: This study aimed to explore the characteristics and the range of evidence on technologies for aging in place by conducting a scoping review of reviews and presenting an evidence map that researchers, policy makers, and practitioners may use to identify gaps and reviews of interest.

Methods: The review was conducted in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews). Literature searches were conducted in Web of Science, PubMed, and Scopus using a search string that consisted of the terms “older people” and “technology for ageing in place,” with alternate terms using Boolean operators and truncation, adapted to the rules for each database.

Results: A total of 5447 studies were screened, with 344 studies included after full-text screening. The number of reviews on this topic has increased dramatically over time, and the literature is scattered across a variety of journals. Vocabularies and approaches used to describe technology, populations, and problems are highly heterogeneous. We have identified 3 principal ways that reviews have dealt with populations, 5 strategies that the reviews draw on to conceptualize technology, and 4 principal types of problems that they have dealt with. These may be understood as methods that can inform future reviews on this topic. The relationships among populations, technologies, and problems studied in the reviews are presented in an evidence map that includes pertinent gaps.

Conclusions: Redundancies and unexploited synergies between bodies of evidence on technology for aging in place are highly likely. These results can be used to decrease this risk if they are used to inform the design of future reviews on this topic. There is a need for an examination of the current state of the art in knowledge on technology for aging in place in low- and middle-income countries, especially in Africa.

(JMIR Aging 2024;7:e50286) doi:10.2196/50286

KEYWORDS
aging in place; technology; gerontechnology; assistive technology; gerontology; geriatric; geriatrics; older adult; older adults; aging; aging; scoping; review methods; review methodology; older people; evidence map; evidence mapping

https://aging.jmir.org/2024/1/e50286
Introduction

Background
The World Health Organization (WHO) estimates that the global population aged 60 years and older will increase from 12% to 22% between 2015 and 2050, with the most dramatic increase in low- and middle-income countries (LMICs) [1]. This will change the age composition in populations globally. Demographic aging refers to shifts in the age composition of populations where the proportion of the population that consists of older people grows significantly. The fact that people are now living longer than ever and that they are expected to continue doing so is the result of positive developments in public health and survival [2]. Yet, demographic aging is also one of the key challenges of our time [3].

This concern is caused by how demographic aging will impact nation states. As people grow older, they tend to become increasingly reliant on both formal and informal care. For instance, older people are more likely to have functional limitations, need assistance with everyday tasks, and need medical care [4-6]. Moreover, older individuals have lower incomes, which compound the challenges of their increasing need for care [7]. As a result, nation states have a variety of systems in place to care for older people, including systems of shouldering the cost of that care.

In countries where welfare and care systems are heavily subsidized, demographic aging is predicted to lead to heavy financial strain and a decreased quality of life for older people, unless solutions that cater to the need to maintain good health and affordable health care into a longer set of retirement years are developed [8]. Still, the adverse consequences of demographic aging will be even greater in LMICs. In LMICs, welfare systems often function poorly or are nonexistent, meaning that the burden of caring for older people falls on families or on the older individuals to care for themselves. This has caused concerns that LMICs will “grow old before growing rich” [7].

To neutralize the overwhelming demand for health care, solutions are needed to enable older people maintain their health for longer and postpone or avoid disability and dependency [1,9,10]. Against this background, interest in technology that enables older people to age in place while maintaining their dignity and quality of life has grown rapidly over the past 2 decades [11].

Aging in place is a concept that refers to the shared responsibility of individuals and public authorities to enable older people to continue to live safely, relatively independently, and comfortably in the community either in their current home or in appropriate housing, regardless of age, income, or level of competence [1,12-14]. The idea is that policies and public services should address the challenges posed by demographic aging by finding alternatives to traditional forms of older adult care and creating solutions that are less resource-intensive. In welfare states where health and care services are heavily subsidized, this shift toward less resource-intensive solutions generally refers to options that maintain a high quality of life for older people while simultaneously preventing or delaying the need to relocate to a nursing home or becoming dependent on care [15], as well as to solutions that minimize the use of resources in nursing homes and other forms of formal care, without compromising their quality. Meanwhile, in LMICs, the main challenge is that welfare systems are weak and even nonexistent. Assistive technology and related services are marginal and not available for the majority, particularly for the poor rural populations. Ensuring assistive technology for all, including the growing number of older adults, requires resources and build-up of competence through a sustainable systems approach [16]. In this context, innovations are needed in service delivery, and community-based models as well as adaptation of existing assistive technology and development of new and contextually relevant assistive technology are needed to ensure that older people live well and as autonomously as possible [17].

Technologies that enable aging in place encompass a wide variety of technologies designed to monitor or support the health and activities of older people or strengthen their contact with others [11,15]. In some cases, older people are the intended users, but technology can also be used to establish links between older adults and their circles of care. Technologies for aging in place include both high- and low-tech solutions, including but not limited to mobility devices, information and communication technologies, assistive technologies, sensor technology, telemedicine, health monitoring, games, wearables and medication reminders, and the internet of things [15,18-27].

Rationale
Alongside the interest in technology that can enable older people to age in place, the number of publications on this topic has increased dramatically. In this context, literature reviews can be important tools for researchers as well as practitioners, policy makers, and decision makers who need to navigate current debates and access syntheses of the available evidence. Yet, to date, there is no review of the available published reviews that provide an overview of the range and characteristics of the evidence on technology for aging in place.

While reviews of reviews on technologies for aging in place do exist, they typically limit the scope to health conditions, diseases, technologies, or caring practices, for instance, by focusing on the self-efficacy of older people using technology to self-manage chronic obstructive pulmonary disease, hypertension, heart failure, or dementia at home [28]; on the effects of digital technologies on older people’s access to health and social care [29]; on the promotion of physical activity in older people using mobile health (mHealth) and eHealth technologies [30]; or on how mHealth technology may support aging in place [31] and procedures of user-centered usability assessment for digital solutions [32].

Objective
The objective of this review of reviews is to explore the characteristics and the range of evidence on technologies for aging in place by conducting a scoping review of reviews in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Systematic Reviews and Meta-Analyses) guidelines.
Scoping Reviews) [33]. The PRISMA-ScR checklist is available in Multimedia Appendix 1.

By exploring the included reviews, we are particularly interested in what year and in which journals they are published, which review methods that characterize reviews in this field, and whether there are any reviews that are explicitly concerned with LMICs. By LMICs, we mean the countries identified by the Organization for Economic Co-operation and Development as having low-income or middle income economies, which may be updated from time to time by the Organization for Economic Co-operation and Development [34]. In exploring the range of evidence presented in reviews on technologies for aging in place, we are particularly interested in which types of populations, technologies, and problems they have been concerned with.

Methods

Eligibility Criteria

We included literature reviews in English about technology for older people or older adult care, including informal care, that we were able to access. To ensure the quality of our sources, we limited our scope to peer-reviewed literature reviews that have been published in academic journals. For the same reason, we only included reviews where the methods were clearly described. We did not apply any limits to the year of publication.

Information Sources

Our method of selecting databases included making a list of the most relevant journals in the field that the authors were aware of (Multimedia Appendix 2). The complete list was sent to a panel of experts consisting of members from the WHO and the International Society of Gerontechnology, who were asked to add any potentially relevant journals missing from the list. After the list was considered complete, the authors identified the databases where these journals were indexed. The final selection of databases was Web of Science (Table 1), PubMed (Table 2), and Scopus (Table 3). The searches in Web of Science and Scopus were conducted on September 13, 2022, and the search in PubMed was conducted on September 14, 2022.

Table 1. Web of Science—core collection (n=1741).

<table>
<thead>
<tr>
<th>Search</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(((((((((TS=(&quot;old* per*&quot;) OR TS=&quot;old* peo*&quot;) OR TS=&quot;old* age*&quot;) OR TS=&quot;old* adu*&quot;) OR TS=&quot;old* use*&quot;) OR TS=(&quot;geriatric&quot;) OR TS=&quot;aged per*&quot;) OR TS=&quot;aged peo*&quot;) OR TS=&quot;aged use&quot;) OR TS=(&quot;ageing&quot;) OR TS=&quot;elder&quot;) OR TS=(&quot;senior&quot;) OR TS=&quot;retire&quot;) OR TS=&quot;pension&quot;) OR TS=&quot;later life&quot;)</td>
<td>4,027,248</td>
</tr>
<tr>
<td>(TS=ai) OR TS=&quot;(ag$ in place&quot;) OR TS=(gerontechnology) OR TS=&quot;(assisted living&quot;) OR TS=&quot;(assist* tech&quot;) OR TS=(assist* device&quot;) OR TS=&quot;tele&quot;) OR TS=&quot;(welfare tech&quot;) OR TS=&quot;(digital* health&quot;) OR TS=&quot;(digital* care&quot;) OR TS=&quot;(smart hom&quot;) OR TS=&quot;(mobile health&quot;) OR TS=(mhealth) OR TS=(ehealth) OR TS=&quot;(robot&quot;)</td>
<td>1,067,363</td>
</tr>
<tr>
<td>#6 AND #5 and Review Article (Document Types) and English (Languages)</td>
<td>1741</td>
</tr>
</tbody>
</table>

Table 2. PubMed (n=2402).

<table>
<thead>
<tr>
<th>Search</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(&quot;old per*&quot;[Title/Abstract] OR &quot;old peo*&quot;[Title/Abstract] OR &quot;old adu*&quot;[Title/Abstract] OR &quot;old use*&quot;[Title/Abstract] OR &quot;geriatric&quot;[Title/Abstract] OR &quot;aged per*&quot;[Title/Abstract] OR &quot;aging&quot;[Title/Abstract] OR &quot;ageing&quot;[Title/Abstract] OR &quot;elder&quot;[Title/Abstract] OR &quot;senior&quot;[Title/Abstract] OR &quot;retire&quot;[Title/Abstract] OR &quot;pension&quot;[Title/Abstract] OR &quot;later life&quot;[Title/Abstract]) AND (english[Filter])</td>
<td>584,813</td>
</tr>
<tr>
<td>(&quot;ai&quot;[Title/Abstract] OR &quot;aging in place&quot;[Title/Abstract] OR &quot;gerontology&quot;[Title/Abstract] OR &quot;assisted living&quot;[Title/Abstract] OR &quot;assistive living&quot;[Title/Abstract] OR &quot;assist* tech&quot;[Title/Abstract] OR &quot;assist* device&quot;[Title/Abstract] OR &quot;tele&quot;[Title/Abstract] OR &quot;welfare tech&quot;[Title/Abstract] OR &quot;digital health&quot;[Title/Abstract] OR &quot;digital care&quot;[Title/Abstract] OR &quot;smart hom&quot;[Title/Abstract] OR &quot;smart hea&quot;[Title/Abstract] OR &quot;mobile health&quot;[Title/Abstract] OR &quot;mhealth&quot;[Title/Abstract] OR &quot;ehealth&quot;[Title/Abstract]) AND (english[Filter])</td>
<td>319,264</td>
</tr>
<tr>
<td>(&quot;independent living&quot;[MeSH Terms] OR &quot;self help devices&quot;[MeSH Terms] OR &quot;artificial intelligence&quot;[MeSH Terms] OR &quot;telemedicine&quot;[MeSH Terms]) AND (review[Filter]) AND (english[Filter])</td>
<td>16,843</td>
</tr>
<tr>
<td>(&quot;aged&quot;[MeSH Terms] OR &quot;aging&quot;[MeSH Terms]) AND (review[Filter]) AND (english[Filter])</td>
<td>11,092</td>
</tr>
<tr>
<td>#1 AND (#2 OR #3 OR #4)</td>
<td>2402</td>
</tr>
</tbody>
</table>
Table 3. SCOPUS (n=3131).

<table>
<thead>
<tr>
<th>#</th>
<th>Search</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TITLE-ABS-KEY ( (&quot;old* per*&quot; ) OR (&quot;old* peo*&quot; ) OR (&quot;old* age*&quot; ) OR (&quot;old* adu*&quot; ) OR (&quot;old* use*&quot; ) OR ( geriatric ) OR (&quot;aged per*&quot; ) OR (&quot;aged peo*&quot; ) OR (&quot;aged use*&quot; ) OR ( aging ) OR ( elder* ) OR ( senior ) OR ( retire* ) OR ( pension* ) OR (&quot;later life&quot; ) )</td>
<td>2,282,529</td>
</tr>
<tr>
<td>2</td>
<td>TITLE-ABS-KEY ( ( ai OR &quot;ag* in place&quot; OR gerontechnology OR &quot;assisted living&quot; OR (&quot;assist* tech&quot; ) OR (&quot;assist* device*&quot; ) OR tele* OR (&quot;welfare tech*&quot; ) OR ( &quot;digital* health&quot; ) OR ( &quot;digital* care&quot; ) OR ( &quot;smart hom*&quot; ) OR ( &quot;smart hea*&quot; ) OR ( &quot;mobile health&quot; ) OR mhealth OR ehealth OR robot* ) )</td>
<td>2,047,409</td>
</tr>
<tr>
<td>3</td>
<td>(LIMIT-TO (DOCTYPE, &quot;re&quot; ) ) AND ( LIMIT-TO ( LANGUAGE , &quot;English&quot; ) )</td>
<td>3131</td>
</tr>
</tbody>
</table>

**Search**

A search consisting of the terms “older people” and “technology for ageing in place” with alternate terms was conducted using Boolean operators and truncation. The search was adapted to the rules for each database.

**Selection of Sources of Evidence**

The search resulted in a total of 7274 identified studies, that is, 3131 from Scopus, 2402 from PubMed, and 1741 from Web of Science. We used Covidence (Veritas Health Innovation) to organize the review process. After 1827 duplicates were identified and removed, 5447 studies were screened using the eligibility criteria (see **Textbox 1**). The original list of eligibility criteria contained items 1-7. However, after we identified a retracted paper, we decided to add exclusion criterion 8 “retracted paper.” The title and abstract screening resulted in the exclusion of 4973 studies. The full-text screening resulted in the further exclusion of 130 studies, and the remaining 344 studies were included in the data charting process. **Figure 1** illustrates this process.

**Textbox 1. Eligibility criteria.**

**Inclusion criteria**
- Literature reviews
- Journal paper
- Peer-reviewed research
- Able to source full text
- Methodologically sound
- About technology for aging in place
- English

**Exclusion criteria, with a short label for Covidence**
- Not a literature review—Papers that do not review the literature
- Not a journal paper—Anything that is not a paper meaning: book chapters, conference proceedings, protocols, reports, preprints, etc
- Not research—editorials, opinion pieces, press, etc
- Unable to source—currently unable to access full text currently
- Method not described—Reviews that do not clearly describe their methods
- Thematically irrelevant—Not about technology for aging in place
- Not in English
- Retracted paper
Data Charting Process and Data Items

The data were extracted using the data extraction template feature in Covidence. The extraction of data was organized in line with our objectives and research questions. Tables 4 and 5 provide an overview of the relationship between the research questions and the extracted data.

Table 4. Research questions and charted data that relate to the characteristics of reviews on technologies for aging in place.

<table>
<thead>
<tr>
<th>Research question</th>
<th>Extracted data</th>
</tr>
</thead>
<tbody>
<tr>
<td>During what years were the reviews published?</td>
<td>Year of publication</td>
</tr>
<tr>
<td>In which journals have the reviews been published?</td>
<td>The name of the journal where they are published</td>
</tr>
<tr>
<td>Which review methods characterize the reviews?</td>
<td>The named review methods they refer to</td>
</tr>
<tr>
<td>Is there an explicit concern with LMICs?(^a)</td>
<td>If they refer explicitly to LMICs (yes or no)</td>
</tr>
</tbody>
</table>

\(^a\)LMIC: low- and middle-income countries.

Table 5. Research questions and charted data that relate to the range of evidence in reviews on technologies for aging in place.

<table>
<thead>
<tr>
<th>Research question</th>
<th>Extracted data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which populations are they concerned with?</td>
<td>The population specified in the review</td>
</tr>
<tr>
<td>Which types of technology are they concerned with?</td>
<td>The technology specified in the review</td>
</tr>
<tr>
<td>What type of problems are they concerned with?</td>
<td>The issues of interest specified in the review</td>
</tr>
<tr>
<td>What is the relationship between the populations, problems, and technologies the reviews have dealt with</td>
<td>The type of technology and the issues of interest specified in the review</td>
</tr>
</tbody>
</table>

All authors participated in the choice of databases and establishing the search terms and eligibility criteria. EL constructed the search string and conducted the final search. All authors participated in the screening process. The full-text papers were extracted by the authors JMB, MG, and AHE. All authors participated in the data synthesis and presentation of the findings.

Results

Characteristics of Evidence on Technologies for Aging in Place

Multimedia Appendix 3 shows an overview of the data and sources that correspond to this section. The number of reviews of evidence on technology for aging in place has increased dramatically over the past few years (Table 6). The earliest review included in our search was published in 2001 [35]. A total of 20 reviews were published between 2001 and 2010. By comparison, 142 reviews were published between 2015 and 2020. Note that the table only includes reviews published before September 13, 2022, when our search was conducted.

From 2020 to 2021, there was a near doubling in number of reviews. Since our search was conducted at the beginning of September 2022, the figure does not show the full extent of published reviews in 2022. However, it is likely that the trend will continue upwards. The included reviews were published in 183 unique journals. Of those, most journals have only published 1 or 2 reviews since 2001. Only 12 journals have published more than 5 reviews in total since 2001 (Table 7).
Table 6. Number of reviews by year of publication (n=344).

<table>
<thead>
<tr>
<th>Year of publication</th>
<th>Reviews, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>2001</td>
<td>1</td>
</tr>
<tr>
<td>2002</td>
<td>0</td>
</tr>
<tr>
<td>2003</td>
<td>0</td>
</tr>
<tr>
<td>2004</td>
<td>1</td>
</tr>
<tr>
<td>2005</td>
<td>1</td>
</tr>
<tr>
<td>2006</td>
<td>0</td>
</tr>
<tr>
<td>2007</td>
<td>5</td>
</tr>
<tr>
<td>2008</td>
<td>4</td>
</tr>
<tr>
<td>2009</td>
<td>6</td>
</tr>
<tr>
<td>2010</td>
<td>2</td>
</tr>
<tr>
<td>2011</td>
<td>4</td>
</tr>
<tr>
<td>2012</td>
<td>8</td>
</tr>
<tr>
<td>2013</td>
<td>12</td>
</tr>
<tr>
<td>2014</td>
<td>14</td>
</tr>
<tr>
<td>2015</td>
<td>6</td>
</tr>
<tr>
<td>2016</td>
<td>16</td>
</tr>
<tr>
<td>2017</td>
<td>23</td>
</tr>
<tr>
<td>2018</td>
<td>23</td>
</tr>
<tr>
<td>2019</td>
<td>33</td>
</tr>
<tr>
<td>2020</td>
<td>47</td>
</tr>
<tr>
<td>2021</td>
<td>88</td>
</tr>
<tr>
<td>2022</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 7. Overview of 12 journals that have published 5 or more reviews on technologies for aging in place since 2001.

<table>
<thead>
<tr>
<th>Journals</th>
<th>Reviews, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Interventions in Aging</td>
<td>5</td>
</tr>
<tr>
<td>Healthcare</td>
<td>5</td>
</tr>
<tr>
<td>Assistive Technology</td>
<td>6</td>
</tr>
<tr>
<td>Journal of Telemedicine and Telecare</td>
<td>7</td>
</tr>
<tr>
<td>Maturitas</td>
<td>8</td>
</tr>
<tr>
<td>JMIR Aging</td>
<td>10</td>
</tr>
<tr>
<td>International Journal of Environment Research and Public Health</td>
<td>11</td>
</tr>
<tr>
<td>Sensors</td>
<td>11</td>
</tr>
<tr>
<td>The Gerontologist</td>
<td>11</td>
</tr>
<tr>
<td>Disability and Rehabilitation: Assistive Technology</td>
<td>13</td>
</tr>
<tr>
<td>International Journal of Medical Informatics</td>
<td>16</td>
</tr>
<tr>
<td>Journal of Medical Internet Research</td>
<td>17</td>
</tr>
</tbody>
</table>

The reviews refer to 15 unique types of review methods. Of these, the most common were systematic reviews (n=144) and scoping reviews (n=60). The third most common review method was to provide a detailed account of the procedures but refrain from referring to a specific type of review method (n=98). While there were only 13 integrative reviews and 6 narrative reviews, the fact that most other review methods only occurred once or twice made the narrative reviews common by comparison (Table 8).
Table 8. Overview of the data analysis methods used in the included reviews.

<table>
<thead>
<tr>
<th>Data analysis method</th>
<th>Reviews, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systematic review</td>
<td>144</td>
</tr>
<tr>
<td>Scoping review</td>
<td>60</td>
</tr>
<tr>
<td>Integrative review</td>
<td>13</td>
</tr>
<tr>
<td>Narrative</td>
<td>6</td>
</tr>
<tr>
<td>Mini-review</td>
<td>5</td>
</tr>
<tr>
<td>Review of reviews</td>
<td>5</td>
</tr>
<tr>
<td>Rapid review</td>
<td>2</td>
</tr>
<tr>
<td>Umbrella review</td>
<td>1</td>
</tr>
<tr>
<td>Targeted review</td>
<td>1</td>
</tr>
<tr>
<td>Meta-interpretive review</td>
<td>1</td>
</tr>
<tr>
<td>Focused literature review</td>
<td>1</td>
</tr>
<tr>
<td>Descriptive review</td>
<td>1</td>
</tr>
<tr>
<td>Clinical review</td>
<td>1</td>
</tr>
<tr>
<td>Critical interpretive synthesis</td>
<td>1</td>
</tr>
<tr>
<td>Conceptual review</td>
<td>1</td>
</tr>
<tr>
<td>Comprehensive review</td>
<td>1</td>
</tr>
<tr>
<td>Comparative literature review</td>
<td>1</td>
</tr>
<tr>
<td>Reflective review</td>
<td>1</td>
</tr>
<tr>
<td>Unspecified</td>
<td>98</td>
</tr>
</tbody>
</table>

Only 1 review referred explicitly to LMICs [36]. This review aimed to identify policy gaps in the delivery and availability of assistive health technology and medical devices for aging populations, particularly in LMICs, and found that practical, life-enhancing support for older people through assistive health technology, medical technology, and related health and social services is a neglected issue.

Range of Evidence on Technologies for Aging in Place Populations

Multimedia Appendix 4 shows an overview of data and sources that correspond to this section. Some reviews dealt with more than one type of population.

Table 9. Overview of the populations in 253 reviews that described the population in terms of people and the roles they play in the context of aging.

<table>
<thead>
<tr>
<th>Populations</th>
<th>Reviews, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Older people</td>
<td>220</td>
</tr>
<tr>
<td>Formal caregivers</td>
<td>1</td>
</tr>
<tr>
<td>Formal caregivers and informal caregivers</td>
<td>2</td>
</tr>
<tr>
<td>Informal caregivers</td>
<td>5</td>
</tr>
<tr>
<td>Older people and formal caregivers</td>
<td>5</td>
</tr>
<tr>
<td>Older people, formal caregivers, and informal caregivers</td>
<td>4</td>
</tr>
<tr>
<td>Older people and informal caregivers</td>
<td>16</td>
</tr>
</tbody>
</table>

In total, 73 reviews described the population in terms of older people as well as individuals from other age groups, with a particular diagnosis or health problem. These reviews included studies about people of different age groups with different cognitive impairments exclusively (n=41) or in combination with other health problems (n=2). Notably then, these reviews...
included evidence based on studies of younger people as well as older people. Table 10 provides an overview of the diagnoses and health problems that these reviews used to conceptualize the populations.

In total, 43 reviews did not specify the population at all. Instead, they referred to the context of aging in place. These reviews were typically concerned with the technical functionality of devices rather than the interplay between what the technology offers and the intended users and their problems.

Table 10. Overview of the diagnoses and health problems used to conceptualize the population in reviews about older people and others with a particular diagnosis or health problem (in total n=73 reviews).

<table>
<thead>
<tr>
<th>Diagnoses and health problems</th>
<th>Reviews, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Cardiovascular diseases</td>
<td>1</td>
</tr>
<tr>
<td>Cardiovascular diseases, diabetes, and asthma</td>
<td>1</td>
</tr>
<tr>
<td>Chronic conditions</td>
<td>7</td>
</tr>
<tr>
<td>Cognitive impairments</td>
<td>41</td>
</tr>
<tr>
<td>Cognitive impairments, cardiovascular diseases, and chronic obstructive pulmonary disease</td>
<td>1</td>
</tr>
<tr>
<td>Cognitive impairments, neurological disorders, falls, and cardiovascular disease</td>
<td>1</td>
</tr>
<tr>
<td>Complex needs</td>
<td>1</td>
</tr>
<tr>
<td>Decline in hand grip and dexterity</td>
<td>1</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1</td>
</tr>
<tr>
<td>Falls</td>
<td>4</td>
</tr>
<tr>
<td>Falls and frailty</td>
<td>1</td>
</tr>
<tr>
<td>Frailty</td>
<td>6</td>
</tr>
<tr>
<td>Frailty and decreased hearing</td>
<td>1</td>
</tr>
<tr>
<td>Hip injuries</td>
<td>1</td>
</tr>
<tr>
<td>Loneliness and social isolation</td>
<td>2</td>
</tr>
<tr>
<td>Mental health</td>
<td>2</td>
</tr>
</tbody>
</table>

Types of Technology

Multimedia Appendix 5 shows an overview of the data and sources that correspond to these results. Some reviews dealt with more than one type of technology.

We identified 69 different types of technology that reviews have been concerned with and 5 substantive strategies that the reviews have used to conceptualize the technology with which they are concerned. Two of the 345 reviews used other strategies for conceptualizing technology. One was about co-designed technologies [37]. The other was about what they termed as consumer technology as well as smart environments [38].

The first and most common strategy (n=140 reviews) is to refer to technology by using descriptive technical terms such as “sensors” [39-48], “artificial intelligence” [49-52], “GPS” [53-55], or “games” [56-60]. The reviews that used this strategy covered 31 different types of technology exclusively or in combination with each other. Most of these reviews were focused on robots or robopets (n=47), information and communication technology (n=23), smart environments (n=17), or sensors (n=10).

The second strategy (n=65) is to conceptualize technology by the purpose of the technology in relation to a disease or type of health challenge that the technology addresses or is believed to be able to address, for instance, by using terms such as “technology for dementia” [61-70], “technologies for social connectedness” [71-75], “technology for frailty” [76-78], “technology for safety” [79], or “technology for falls” [60,80-84]. Most of the reviews that relied on this strategy dealt with assistive technologies (n=28).

The third strategy (n=30 reviews) is to refer to technologies in terms of their intended purpose in caring services or practices that they are part of. Such terms include “teleophthalmology” [85], “monitoring technologies” [86,87], “telerehabilitation” [88-93], “technology for home health care” [94,95], or “technology for pain management” [96]. Most reviews that relied on this strategy were concerned with telerehabilitation (n=7) or technology for health information (n=6).

The fourth strategy (n=29 reviews) is to describe the type of technology by using umbrella terms that broadly refer to the use of technology to enable older people to age in place, for instance, by defining the technology of interest in terms of “technology for ageing in place” [11,97-99], “gerontechnology” [100-102], “welfare technology” [103-105], “technology for healthy ageing” [106,107], or “technology for older people” [22,32,108-112].

The fifth strategy (n=75 reviews) is to describe the type of technology the review is concerned with by way of concepts that refer to the use of technology as part of a broad range of caring services, strategies, and practices, such as, for instance,
telecare [113-119], telemedicine [120-125], e-interventions [126], or eHealth [127-143]. Most reviews that used this type of concept to describe the technology they are concerned with dealt with mHealth (n=18), eHealth (n=17), or telehealth (n=18).

What Types of Problems Have the Reviews Dealt With? Multimedia Appendix 6 shows an overview of the corresponding data and sources. Some reviews dealt with more than one type of problem. We identified 49 unique problem topics and 4 principal types of problems.

The first type of problem is related to different types of care services or caring practices (n=60 reviews). Most of these reviews dealt with problems related to the context of home care (n=30), caring practices in nursing homes or other long-term care institutions (n=11), or rehabilitation (n=7). By contrast, other topics occurred only once or twice, that is, problems related to caregiver burden [144,145], dementia care [146], emergency care services [147], informal care [148], and health information services [149,150].

The second type of problem is issues related to the management of health-related issues or diseases in the context of aging in place (n=128 reviews). Of those, most dealt with problems related to cognitive impairments either exclusively (n=61) or in combination with one or several other health problems (n=10), that is, cognitive impairment and mental health [151-155], or cognitive impairment, stroke, cardiovascular disease, and falls [156]. Other problems that were featured relatively frequently included falls and balance–related issues (n=19), frailty (n=8), chronic conditions (n=8), and depression (n=5). Meanwhile, other problems related to the management of other health-related issues and diseases featured only once or twice, despite being common health challenges for older people (for instance, Parkinson’s disease [157,158], malnutrition [159], dental health [160], eye diseases [85], and pain [96,161]).

The third type of problem relates to the experience of aging in place (n=82 reviews). The most common topics in this category were loneliness, including social isolation or connectedness (n=21 reviews), older peoples’ self-care or self-management (n=19 reviews), and active aging (n=16 reviews). Other topics in this category include healthy aging [106,111,137,162-165], information needs [166], quality of life [97,167-173], quality of life and older people’s self-care and self-management [174,175], and activities of daily living exclusively [176-179] or in combination with other topics such as loneliness [180,181], or quality of life [182].

The fourth type of problem relates to the research and development of technology. This was the most common type of problem (n=285 reviews). The overwhelming majority of reviews that dealt with this type of problem were concerned with barriers and drivers of use and acceptability (n=114), the effect or implications of technology (n=86), or the combination of these 2 topics (n=21). Other common topics included uptake or scalability (n=15), user involvement (n=11), ethical considerations (n=14), feasibility (n=10), and cost-effectiveness or use (n=7).

Notably, problems related to home care (n=30), loneliness (n=21), cognitive impairments (n=71), barriers and drivers of use and acceptability (n=114 reviews), and the effect or implications of technology (n=86) have been heavily emphasized. Meanwhile, others such as cost-effectiveness or use of technologies (n=7), health information needs (n=1), malnutrition (n=1), dental health (n=1), eye diseases (n=1), and pain management seem underprioritized by comparison.

What Are the Relationships Between the Problems, Technologies, and Populations That the Reviews Have Dealt With

Multimedia Appendix 7 shows an evidence map that provides an overview of the relationships between problems, technologies, and populations that the reviews have been concerned with. Some reviews deal with more than one population, technology, and type of problem. Multimedia Appendix 8 shows an overview of the corresponding data and sources.

As illustrated in the evidence map (Multimedia Appendix 7), many reviews draw on an evidence base that is not specific to older people or their caregivers.

This is particularly notable in the reviews on the following topics: barriers and drivers of use and acceptability, cognitive impairment, and the effect or implications of technology. The same observation applies to the following types of technology such as assistive technologies, robots, technology for dementia, technology for falls, technology for frailty, telehealth, and technology for Alzheimer disease.

Summary of Evidence

In exploring the range and characteristics of reviews on technology for aging in place, we found that the number of reviews, as well as the pace at which they are published, has increased dramatically over time. While some journals such as JMIR Aging, Disability and Rehabilitation: Assistive Technology, the Journal of Medical Internet Research, and The Journal of Medical Informatics have published more reviews on this topic than others, the literature is scattered over 183 unique journals. Most reviews on this topic are systematic reviews (n=144).

In exploring the range of reviews on technology for aging in place, we identified 3 principal ways that reviews have dealt with populations. Specifically, the 3 ways are describing the population in terms of older people or different types of caring roles (n=253), in terms of people affected with a particular health condition or diagnosis (n=73), or not specifying the population (n=43). These may be considered as methods of conceptualizing populations. We identified 88 unique types of technology that the reviews have dealt with. We also found that there are strong tendencies for reviews to synthesize the evidence on broad and unspecific categories of technology such as “ICT” or “robots” rather than to concentrate on a particular device (a notable exception is a review on personal alarms [183]). Moreover, we identified 5 strategies that the reviews draw on to conceptualize technology. Those strategies are to (1) refer to technology by using descriptive technical terms; (2) conceptualize technology by way of the purpose of the technology about a disease or health issue; (3) refer to technologies in terms of their purpose in caring services or
practices; (4) use umbrella terms that broadly refer to the use of technology to enable older people to age in place; and (5) use concepts that refer to the technology as part of caring services, strategies, and practices. We also identified 4 principal types of problems and 49 unique subtypes of problems that the reviews have dealt with. The four principal types are problems related to (1) different types of care services or caring practices, (2) the management of health problems or diseases, (3) the experience of aging in place, and (4) the research and development of technology. The evidence map (Multimedia Appendix 6) demonstrates the relationships between the populations, technologies, and problems studied in the reviews and illustrates the gaps. Notably, many of the reviews on the most studied technologies and problems draw on studies that are not specific to older people or the context of aging in place, either by not specifying the population at all or by including studies on patients of all ages, meaning that topics studied only by such reviews should also be considered gaps.

**Discussion**

**Summary of Evidence**

Together, these results speak to the need for regularly updated overviews of ongoing debates in the field. However, they are also illustrative of the challenges that such overviews must overcome. For instance, the lack of conceptual hegemony means that any attempt to describe the technologies that the reviews have been concerned with in purely technical terms fails to grasp the diverse ways that technology is understood in this field. A more fruitful approach is to categorize them according to the different ways that they understand and deal with technology. Used as methodological tools, the strategies of defining populations, conceptualizing technology, the typology of problems, and the overview of the relationships presented here can inform the design of future reviews and enable researchers to purposefully identify gaps and publications that are likely to be of relevance to each other despite conceptual differences that may obscure their similarities.

It is notable that only 1 review was explicitly concerned with LMICs, considering that the largest growth in older people globally will be in LMICs [1], particularly in Africa where the population of 60 years and older is expected to increase by more than 100% by 2050 [184]. Similarly, it is notable that in the included reviews, relatively little attention has been paid to formal and informal caregivers. Both formal and informal caregivers play important roles in the context of technology for aging in place. Both formal and informal caregivers frequently speak and act on behalf of older people, especially older people with cognitive impairments when technology developers seek to identify user needs or evaluate the usefulness of the technology [185-188]. In doing so, they act as gatekeepers who shape what types of technology are developed and offered to older people, and equally important, which are not [109]. Both formal and informal caregivers are often the intended users of technology that is meant to enable older people to age in place. Thus, the politics of their lives and working conditions as well as the quality and type of care they are able to provide to older people are shaped by what the technology affords and prohibits [189-191]. Yet, the purpose of the technology is aimed at the needs of the older person or efficiency-related goals in care organizations rather than the improvement of the care workers’ working environment or care burden. Additionally, like all users, both formal and informal caregivers are not just impacted by technologies that enter their lives but they also shape the technology in turn [188,192-199], meaning that the implications that the technology will have in practice are never given beforehand and must always be studied in the context of use [185,187,200,201]. Finally, both informal and formal caregivers must frequently improvise and adapt the technology to render it functional [192,202-205]. Thus, both formal and informal caregivers play important roles in shaping the practices, politics, and services that the technology affords or delimits in the lives of older people who age in place. These roles have been thoroughly described in the literature. Yet, they seem overlooked in reviews on technology for aging in place.

It is problematic that so many reviews concerned with problems related to technologies for aging in place draw on an evidence base that is not specific to older people. Older people frequently have other needs than younger people even when they share a diagnosis because the aging body presents specific challenges, which increase the risk of illnesses, falls, disability, and death [206]. It is therefore unlikely that reviews that do not focus explicitly on older people are able to grasp and address the specificity of the challenges that older people face as part of aging in place. This primarily concerns reviews on the topics of barriers and drivers of use and acceptability, cognitive impairment, and the effect or implications of technology. It also concerns reviews about assistive technologies, robots, technology for dementia, technology for falls, technology for frailty, telehealth, and technology for Alzheimers disease. While these topics and technologies have frequently been addressed, the value that reviews that do not specify their population or that base their arguments on studies of people of all ages (see Multimedia Appendix 6) is limited, and there is a need for more targeted and age-specific syntheses reviews to better address the unique requirements of older individuals and their caregivers. The strong tendency for reviews in this field to concentrate on broad and unspecific categories of technology, such as “ICT” or “robots” means that there is no straightforward way for practitioners to use these reviews as support in decision-making processes regarding the potential usefulness and challenges related to specific devices.

**Limitations**

Despite the many methodological strengths of the design of a scoping review of reviews, there are some limitations to be considered. These include the potential for bias in the review process, the difficulty ensuring the quality and reliability of the included reviews, and the potential for the review to be influenced by the perspectives and priorities of the researchers conducting the review. Considering the broad eligibility criteria chosen for this review, the results may be considered representative of the characteristics and range of evidence on technologies for aging in place. However, the inclusion of more databases could have expanded the data set even further, and potentially relevant literature that does not use the term aging in place explicitly may have been missed. Moreover, this review
has not sought to explore or synthesize the results of the included reviews nor have we considered the quality of the included reviews.

Conclusions
The number of published reviews on this topic in the past few years in combination with the rate at which they are published suggests that redundancies and a lack of fruitful synergies between them are likely. The breadth of variation concerning how reviews have dealt with populations, conceptualizations of types of technology, and problems demonstrates the conceptual differences that must be bridged to remedy this problem.

Together, these results underscore the necessity for improved coordination and collaboration among reviews while also recognizing the potential benefits of more standardized vocabularies.

The insights gained from the methods of dealing with populations, strategies for conceptualizing types of technology, and the types of problems identified in this study may be used methodologically to identify commonalities and connections that may otherwise be obscured by differing conceptual frameworks.

There is an urgent need for an examination of the current state of the art in knowledge regarding technology for aging in place in LMICs. Developing a deeper understanding of the conditions surrounding aging in LMICs, especially in Africa, and the implications those conditions have for the roles that technology may play and not play in the lives of older people and their circles of care should be an essential focus of the research agenda.

Acknowledgments
The scoping review was conducted by SINTEF as part of a collaboration among SINTEF, WHO, and the International Society for Gerontechnology (ISG) that aim to ensure an emphasis on older people within the current global momentum on AT. The authors would like to thank Espen H. Aspnes, former vice president of ISG who leads Pillar II in this collaboration, to which the result of this review contributes. We also wish to thank Professor Sue Levkoff (ScD, MSW, SM), College of Social Work, the University of South Carolina. Endowed chair in SeniorSMART, as well as other members of the International Society for Gerontechnology (ISG) and The WHO contributed to discussions and workshops during the planning stage of this review. For their contributions to the selection of databases, we especially thank Dr. Callista Kahonde, researcher and assistant lecturer at Stellenbosch University; Centre for Disability and Rehabilitation Studies, and Anna Spånt Enebuske, MSSc and research officer at The Swedish Municipal Workers’ Union. Finally, we thank our colleague Sandra Klonteyer MSc, SINTEF Digital, for her contributions during the data extraction phase, and Mr. Sebastian M Bergschöld for his contributions to the evidence map. This study was funded by SINTEF.

Conflicts of Interest
None declared.

Multimedia Appendix 1
PRISMA-ScR Checklist.
[DOCX File, 56 KB - aging_v7i1e50286_app1.docx]

Multimedia Appendix 2
Journals to guide the selection of databases.
[PDF File (Adobe PDF File), 116 KB - aging_v7i1e50286_app2.pdf]

Multimedia Appendix 3
Characteristics of evidence on technologies for ageing in place.
[XLSX File (Microsoft Excel File), 68 KB - aging_v7i1e50286_app3.xlsx]

Multimedia Appendix 4
Populations.
[XLSX File (Microsoft Excel File), 62 KB - aging_v7i1e50286_app4.xlsx]

Multimedia Appendix 5
Types of technology.
[XLSX File (Microsoft Excel File), 69 KB - aging_v7i1e50286_app5.xlsx]

Multimedia Appendix 6
Types of problems.
[XLSX File (Microsoft Excel File), 66 KB - aging_v7i1e50286_app6.xlsx]

Multimedia Appendix 7
Evidence map.
[XLSX File (Microsoft Excel File), 71 KB - aging_v7i1e50286_app7.xlsx]

Multimedia Appendix 8
Relationships between problems technologies and populations.
[XLSX File (Microsoft Excel File), 75 KB - aging_v7i1e50286_app8.xlsx]

References


Sülö, S; van Elten, HJ; Askari, M; Weggelaar-Jansen, AM; Huijsman, R. eHealth applications to support independent living of older persons: scoping review of costs and benefits identified in economic evaluations. J Med Internet Res 2021;23(3):e24363 [FREE Full text] [doi: 10.2196/24363] [Medline: 33687335]


205. Styhre A. Tinkering with material resources: operating under ambiguous conditions in rock construction work. Learn Organ 2009;16(5):386-397 [FREE Full text] [doi: 10.1108/09696470910974171]


Abbreviations

- **LMIC**: low- and middle-income country
- **mHealth**: mobile health
- **PRISMA-ScR**: Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Review
- **WHO**: World Health Organization

© Jenny M Bergschöld, Mari Gunnes, Arne H Eide, Eva Lassemo. Originally published in JMIR Aging (https://aging.jmir.org), 22.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
The Prevalence of Missing Incidents and Their Antecedents Among Older Adult MedicAlert Subscribers: Retrospective Descriptive Study

Antonio Miguel-Cruz¹*, PhD; Hector Perez²*, PhD; Yoojin Choi³*, PhD; Emily Rutledge², BA; Christine Daum²*, PhD; Lili Liu²*, PhD

¹Department of Occupational Therapy, Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, AB, Canada
²Faculty of Health, University of Waterloo, Waterloo, ON, Canada
³Faculty of Medicine, University of British Columbia, Vancouver, BC, Canada
*these authors contributed equally

Corresponding Author:
Antonio Miguel-Cruz, PhD
Department of Occupational Therapy
Faculty of Rehabilitation Medicine
University of Alberta
116 St & 85 Ave
Edmonton, AB, T6G 2R3
Canada
Phone: 1 7802246641
Email: miguelcr@ualberta.ca

Abstract

Background: With the population aging, the number of people living with dementia is expected to rise, which, in turn, is expected to lead to an increase in the prevalence of missing incidents due to critical wandering. However, the estimated prevalence of missing incidents due to dementia is inconclusive in some jurisdictions and overlooked in others.

Objective: The aims of the study were to examine (1) the demographic, psychopathological, and environmental antecedents to missing incidents due to critical wandering among older adult MedicAlert Foundation Canada (hereinafter MedicAlert) subscribers; and (2) the characteristics and outcomes of the missing incidents.

Methods: This study used a retrospective descriptive design. The sample included 434 older adult MedicAlert subscribers involved in 560 missing incidents between January 2015 and July 2021.

Results: The sample was overrepresented by White older adults (329/425, 77.4%). MedicAlert subscribers reported missing were mostly female older adults (230/431, 53.4%), living in urban areas with at least 1 family member (277/433, 63.8%). Most of the MedicAlert subscribers (345/434, 79.5%) self-reported living with dementia. MedicAlert subscribers went missing most frequently from their private homes in the community (96/143, 67.1%), traveling on foot (248/270, 91.9%) and public transport (12/270, 4.4%), during the afternoon (262/560, 46.8%) and evening (174/560, 31.1%). Most were located by first responders (232/486, 47.7%) or Good Samaritans (224/486, 46.1%). Of the 560 missing incidents, 126 (22.5%) were repeated missing incidents. The mean time between missing incidents was 11 (SD 10.8) months. Finally, the majority of MedicAlert subscribers were returned home safely (453/500, 90.6%); and reports of harm, injuries (46/500, 9.2%), and death (1/500, 0.2%) were very low.

Conclusions: This study provides the prevalence of missing incidents from 1 database source. The low frequency of missing incidents may not represent populations that are not White. Despite the low number of missing incidents, the 0.2% (1/500) of cases resulting in injuries or death are devastating experiences that may be mitigated through prevention strategies.

(JMIR Aging 2024;7:e58205) doi:10.2196/58205

KEYWORDS
dementia; Alzheimer disease; memory loss; cognitive impairment; missing incident; wandering; critical wandering; older adults; retrospective design
Introduction

Background

With the population aging, the number of people living with dementia will increase. Currently, approximately 55 million people have dementia globally [1]. With approximately 10 million new diagnoses each year, the total number of cases is expected to rise to 78 million by 2030 [1]. Canada is no exception to this trend, with at least 546,000 people currently living with dementia [2]. By 2030, the number of Canadians with dementia will reach at least 1,712,400 [3].

The disease burden cost associated with dementia is sizable in Canada. It was estimated that the direct costs (eg, long-term care) associated with dementia was CAD $10.4 billion (US $7.52 billion) in 2016, and it is expected to double by 2030 [2]. Half of the global cost of dementia is attributed to informal care (ie, family members and friends) [1]. It is estimated that, on average, care partners spend 26 to 35 hours per week caring for persons with dementia [1,4]. This overwhelming number of caregiving hours is attributed to personal care (ie, personal care such as bathing, feeding, and assisting with toileting) [5] and vigilance as a prevention strategy to prevent unattended exits, ultimately keeping persons living with dementia safe in their homes [6].

With increasing numbers of people living with dementia, the prevalence of missing incidents due to critical wandering is rising as well (refer to section The Concept of Missingness and Critical Wandering and Its Risk Factors: A Brief Theoretical Background). However, research on prevalence estimation on missing incidents due to critical wandering is inconclusive in some jurisdictions and overlooked in others. Limitations of the prior literature on this topic exist. First, there is a lack of consistency on reported prevalence [7], leading to disparate statistics; for example, McShane et al [8] reported that 40% of people with dementia become lost, and 5% do so repeatedly. The Alzheimer’s Association estimated that 60% of people with Alzheimer disease will wander [3], and a considerably larger set of studies showed that the prevalence of wandering varies from 11% to 60% [9,10]. Second, previous studies included low sample sizes from limited secondary data sources (eg, data not retained for >5 years) [11] such as police data and data obtained from newspaper report or social media [12-15], leading to a limited scope of the statistical analyses. Third, prevalence studies have been completed in the United States [16,17], Japan [18-20], and South Korea [21,22], leaving the prevalence of missing incidents among people with dementia in Canada largely unknown. This is an important gap in our knowledge because Canada has distinct social, health care, and geographic features as well as a harsh climate, making it challenging to extrapolate data from other countries for its unique context. As a result, the prevalence and risk factors of missing incidents due to critical wandering for Canadians living with dementia remain largely unknown.

Objectives

The aims of this study were to examine (1) the demographic, psychopathological, and environmental antecedents to missing incidents due to critical wandering among older adult MedicAlert subscribers; and (2) the characteristics and outcomes of the missing incidents. We used a retrospective descriptive design. The sample included 434 older adult MedicAlert subscribers (hereinafter MedicAlert subscribers) involved in 560 missing incidents between January 2015 and July 2021.

The Concept of Missingness and Critical Wandering and Its Risk Factors: A Brief Theoretical Background

Missingness is the phenomenon of going missing [11]. A missing person is an “individual that cannot be found” [23]. A missing person is “an individual whose whereabouts are unknown to members of their familial, social or professional networks where there is concern for either their own safety and wellbeing or that of others” [24]. A person can go missing intentionally or unintentionally. A person who goes missing unintentionally is said to be lost if the person is (1) “confused with current location in respect to finding other locations” and (2) “unable to reorient” [23]. In this research, we analyzed missing incidents related to persons (older adults) who go missing unintentionally. People living with dementia are at risk of unintentionally getting lost due to critical wandering. “Critical wandering” occurs when an individual living with dementia “leaves an institution or home [with or without the consent of their care partner] and is unaware of his or her situation in terms of place and/or time” [7]; the person is lost. Critical wandering is a necessary (but not a sufficient) condition for a missing incident to occur. A missing incident of a person living with dementia can occur when, for instance, this person is left unsupervised for a few minutes and has an episode of critical wandering [25]. Therefore, critical wandering and missing incident are 2 distinct concepts, although the literature in this field acknowledges that the former could lead to the latter [26].

Antecedents or risk factors influence the mechanisms preceding and contributing to missing incidents [11,26,27]. One way to classify antecedents is to determine whether characteristics are intrinsic (demographic and psychopathological or neurocognitive antecedents) or extrinsic (situational or environmental antecedents) to the missing individual. Demographic antecedents comprise sex, gender, age, and ethnicity of the missing individual. Psychopathological or neurocognitive antecedents are manifestations of behaviors related to cognitive or psychological impairment or mental illness, disorders, or distress. Finally, environmental antecedents may include social, cultural, political, economic, and weather conditions [11]. Another way to classify antecedents is to determine whether they are fixed or variable. A fixed antecedent is one that does not vary within individuals over time (eg, ethnicity). Conversely, a variable antecedent changes over time (eg, the age of an individual) [28]. Missing incidents can lead to consequences or outcomes for the missing person and their care partners [26]. For a missing person, these outcomes can range from returning home safely to minor injuries, major injuries, or even death [26].

MedicAlert Service

MedicAlert is a Canadian-based service that can assist first responders and Good Samaritans in identifying an individual who has gone missing and connecting them with their care partners to help them return to their place of residence. The 2
primary tools used in the service are a medical ID object and a personal health information record. The ID object, which is typically in the form of a bracelet, contains key health conditions and a unique MedicAlert ID number specific to the individual. The MedicAlert ID number can then be used by authorized personnel to access a subscriber’s personal health information record, which contains extensive details about the subscriber’s medical conditions and medications as well as information on previous wandering history if provided by the subscriber or care partner. It is important to recognize that information is self-reported instead of provided by, or confirmed with, health providers, that is, a MedicAlert subscriber or care partners are at liberty to disclose details about the person’s situation and medical condition when the person goes missing. This information is relayed via a 24/7 hotline or by direct digital access by police dispatchers. When the missing person is found, the hotline operator or the police dispatcher notifies the care partners about the missing person’s location [29]. By linking care partners, first responders, and Good Samaritans, the goal is to safely return the missing person back home.

Textbox 1. Study variables.

<table>
<thead>
<tr>
<th>Variable and description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographic antecedents: age, sex at birth, ethnic background, official Canadian languages spoken, province, and primary contact</td>
</tr>
<tr>
<td>Psychopathological antecedents: medical conditions</td>
</tr>
<tr>
<td>Environmental antecedents: domicile (urban vs rural) and living arrangement</td>
</tr>
</tbody>
</table>

Characteristics of the missing incidents: mode of mobility, the time of day and season in which the missing incident occurred, the family care partner’s involvement in response to the missing incident, who reported and found the missing person, point last seen or where the person was missing from, location in which the person was found (actual and self-reported), number of missing incidents, repeated missing incident history (actual and self-reported), mean time to the first missing incident (in days), mean time between missing incidents (in days), and survivability

Procedures

**Missing Incident Selection Procedure**

Detailed information about the missing incidents were obtained from the summary notes made by the MedicAlert hotline operator when a call was received at the MedicAlert call center. These notes are documented by the operator each time a call is received and were in free-text format. We included missing incidents in which the MedicAlert subscribers (1) were aged ≥65 years, (2) went missing unintentionally, and (3) there was clear indication that the subscriber was actually lost (indications of disorientation or confusion or spatial navigation challenges). We excluded missing incidents that (1) were false positive reports (eg, GPS devices were activated and generated a record in the hotline access database, miscommunication between family members, and missing incident calls created for training purposes), (2) were a duplicate missing incident in which several follow-up calls were associated with the same missing incident, or (3) did not contain enough information from which to extract data.

Upon receiving the data set, 8 team members (including the authors) immersed themselves in the data set. Each read 60 different call summary notes and made notes on their contents in relation to the free-text fields in the MedicAlert subscriber database (eg, domicile [urban vs rural]; refer to Multimedia Appendix 1 for more details). Team members shared their observations during 2 subsequent meetings, and these were used to create a preliminary coding framework, including key definitions and the operationalization of each variable (refer to the next subsection for more details). Two team members, hereinafter referred to as analysts, screened, extracted, and coded relevant information from the summary notes made by the MedicAlert hotline operator using the coding framework. The analysts then completed a calibration exercise in which they independently applied the inclusion and exclusion criteria as well as coded 10 included missing incidents. The calibration exercise was conducted as follows: first, 10 cases were selected randomly; next, 2 researchers independently assessed the cases based on the inclusion and exclusion criteria; and, finally, the 2 researchers met and debriefed on the main causes of disagreements. In the calibration process, the team achieved a 90% agreement (ie, percentage of agreement calculated as the number of times a set of ratings is the same divided by the total number of units of observation that are rated). The analysts then screened and coded data for another 100 missing incidents (registries) independently and checked each other’s work. The coding framework was revised to improve the clarity of the definitions and the operationalization of the codes. The analysts met weekly to discuss missing incidents that were unclear or required a second opinion and revised the coding framework to increase clarity. When the coding framework was revised, the analysts reviewed the previously coded data against the revised
coding framework and recoded as necessary. The analysts also sought feedback from the first author when conflicts arose in their screening and coding. In total, 7045 missing incidents were screened from the hotline access MedicAlert database; after applying the inclusion and exclusion criteria, 6485 (92.5%) incidents were excluded. The 6485 missing incidents were excluded due to false positive reports (n=5093, 78.84%), not enough information from which to extract data (n=1076, 16.66%), no indications of disorientation or confusion or spatial navigation challenges (n=270, 3.8%), and MedicAlert subscribers being aged <65 years (n=46, 0.65%).

Categories Generation and Operationalization of Variables

After the data set was cleaned, variables regarding antecedents to the missing incident were coded categorically based on previous research [30,31] and following Statistics Canada classifications whenever possible [32]. As some variables were stored in the form of free text, categories were generated inductively from the information contained in the free text (refer to the preceding subsection; eg, missing incident notes compiled by the hotline operator). Finally, all variables were operationalized as follows: dichotomous variables were coded as 0 or 1 (eg, MedicAlert subscriber’s sex), and each polytomous variable was represented by a set of binary variables, whose values codified each variable category.

Data Analyses

We used descriptive statistics, including mean and SDs, to summarize the central values of distributions for continuous variables. We used the chi-square and Fisher exact—in the case of small, expected counts—tests for comparing categorical variables. Where appropriate, t tests (2-tailed) or the Mann-Whitney U test (2 independent groups, 2-tailed) and 1-way ANOVA (2-tailed) or the Kruskal-Wallis rank sum test (>2 groups, 2-tailed) were used for determining the difference between groups for continuous variables. Where appropriate, we used Cramer V and Pearson and Spearman ρ to determine correlations or associations between variables. As this was a retrospective descriptive study, each variable was examined separately [31]. Statistical analysis was conducted using SPSS software (version 28.0; IBM Corp). The α was set at .05.

Results

Demographic, Psychopathological, and Environmental Antecedents to Missing Incidents Among MedicAlert Subscribers

Table 1 shows the demographic and environmental antecedents to missing incidents among MedicAlert subscribers. Overall, 434 MedicAlert subscribers were involved in 560 missing incidents between January 2015 and July 2021. Regarding psychopathological or neurocognitive antecedents, in 79.5% (345/434) of the cases, MedicAlert subscribers self-reported living with dementia, and the remaining 20.5% (89/434) self-reported having other medical conditions, the most prevalent being short- and long-term memory loss and mental health issues such as depression, schizophrenia, and anxiety disorder. However, it is important to keep in mind that these data are self-disclosed at the time of subscribing to MedicAlert and thus may underestimate the true prevalence of dementia in this sample. The average age of the MedicAlert subscribers was 82.56 (SD 7.4) years, with a little more than half (230/431, 53.4%) identifying as female. The most prevalent age groups were 75 to 84 years (177/434, 40.8%) and 85 to 94 years (168/434, 38.7%), together representing 79.5% (345/434) of the sample. White older adults represented the vast majority (329/425, 77.4%) of the subscribers. In 55.8% (240/430) of the cases, the subscribers spoke English, with an additional 18.1% (78/430) who spoke another language or other languages in addition to English; notably, 11.6% (50/430) of the subscribers spoke neither of the 2 official Canadian languages, English and French. MedicAlert subscribers primarily resided in Ontario (199/341, 58.3%), British Columbia (57/341, 16.7%), or Quebec (50/341, 14.7%); and a vast majority (331/341, 97.1%) lived in urban areas. Living arrangements included with family (277/433, 64%) and in a facility (90/433, 20.8%), although 13.1% (57/433) reported living alone. Most of the subscribers (309/341, 90.6%) listed family members as their primary contact.
Table 1. Demographics and environmental antecedents of the sample (unit of analysis: MedicAlert subscriber).

<table>
<thead>
<tr>
<th>Demographic characteristics</th>
<th>Persons without dementia and persons living with dementia</th>
<th>Persons without dementia</th>
<th>Persons living with dementia</th>
<th>Statistical tests (persons without dementia vs persons living with dementia)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F test (df)</td>
<td>Chi-square (df)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (y), mean (SD; range)</td>
<td>82.56 (7.4; 65-101)</td>
<td>83.90 (7.153; 66-101)</td>
<td>82.21 (7.381; 65-99)</td>
<td>0.93 (1,429)</td>
<td></td>
</tr>
<tr>
<td>Sex assigned at birth (n=431), n (%)</td>
<td>—</td>
<td>0.9 (1)</td>
<td>0.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>230 (53.4)</td>
<td>51 (11.8)</td>
<td>179 (41.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>201 (46.6)</td>
<td>37 (8.6)</td>
<td>164 (38.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age group (y; n=434), n (%)</td>
<td>—</td>
<td>7.8 (3)</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65-74</td>
<td>72 (16.6)</td>
<td>9 (2.1)</td>
<td>63 (14.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75-84</td>
<td>177 (40.8)</td>
<td>35 (8.1)</td>
<td>142 (32.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>85-94</td>
<td>168 (38.7)</td>
<td>38 (8.8)</td>
<td>130 (29.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95-104</td>
<td>17 (3.9)</td>
<td>7 (1.6)</td>
<td>10 (2.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;105</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnic background (n=425), n (%)</td>
<td>—</td>
<td>3.4 (3)</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>23 (5.4)</td>
<td>6 (1.4)</td>
<td>17 (4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chinese</td>
<td>20 (4.7)</td>
<td>5 (1.2)</td>
<td>15 (3.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>329 (77.4)</td>
<td>67 (15.8)</td>
<td>262 (61.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>53 (14.3)</td>
<td>6 (1.4)</td>
<td>47 (11.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Official Canadian languages spoken (n=430), n (%)</td>
<td>—</td>
<td>6.4 (5)</td>
<td>0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English only</td>
<td>240 (55.8)</td>
<td>47 (10.9)</td>
<td>193 (44.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English and other</td>
<td>78 (18.1)</td>
<td>11 (2.6)</td>
<td>67 (15.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neither English nor French</td>
<td>50 (11.6)</td>
<td>10 (2.3)</td>
<td>40 (9.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>French only</td>
<td>30 (7)</td>
<td>10 (2.3)</td>
<td>20 (4.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>French and English</td>
<td>27 (6.3)</td>
<td>6 (1.4)</td>
<td>21 (4.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>French and other</td>
<td>5 (1.2)</td>
<td>2 (0.5)</td>
<td>3 (0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Province (n=341), n (%)</td>
<td>—</td>
<td>6.6 (7)</td>
<td>0.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ontario</td>
<td>199 (58.4)</td>
<td>36 (10.6)</td>
<td>163 (47.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>British Columbia</td>
<td>57 (16.7)</td>
<td>11 (3.2)</td>
<td>46 (13.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quebec</td>
<td>50 (14.7)</td>
<td>16 (4.7)</td>
<td>34 (9.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alberta</td>
<td>18 (5.3)</td>
<td>3 (0.9)</td>
<td>15 (4.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manitoba</td>
<td>9 (2.6)</td>
<td>2 (0.6)</td>
<td>7 (2.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nova Scotia</td>
<td>4 (1.2)</td>
<td>0 (0)</td>
<td>4 (1.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saskatchewan</td>
<td>3 (0.9)</td>
<td>1 (0.3)</td>
<td>2 (0.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Brunswick</td>
<td>1 (0.3)</td>
<td>0 (0)</td>
<td>1 (0.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Newfoundland and Labrador</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prince Edward Island</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nunavut and Northwest Territories</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yukon</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population density (n=341), n (%)</td>
<td>—</td>
<td>0.6 (1)</td>
<td>0.44</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

https://aging.jmir.org/2024/1/e58205

JMIR Aging 2024 | vol. 7 | e58205 | p.243
(page number not for citation purposes)
### Demographic characteristics

<table>
<thead>
<tr>
<th>Demographic characteristics</th>
<th>Persons without dementia and persons living with dementia</th>
<th>Persons without dementia</th>
<th>Persons living with dementia</th>
<th>Statistical tests (persons without dementia vs persons living with dementia)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F$ test ($df$)</td>
<td>Chi-square ($df$)</td>
<td>$P$ value</td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>331 (97.1)</td>
<td>66 (19.4)</td>
<td>265 (77.7)</td>
<td></td>
</tr>
<tr>
<td>Rural</td>
<td>10 (2.9)</td>
<td>3 (0.9)</td>
<td>7 (2.1)</td>
<td></td>
</tr>
<tr>
<td><strong>Living arrangement (n=433), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>With family</td>
<td>277 (64)</td>
<td>53 (12.2)</td>
<td>224 (51.7)</td>
<td></td>
</tr>
<tr>
<td>Facility</td>
<td>90 (20.8)</td>
<td>24 (5.5)</td>
<td>66 (15.2)</td>
<td></td>
</tr>
<tr>
<td>Alone</td>
<td>57 (13.1)</td>
<td>9 (2.1)</td>
<td>48 (11.1)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>9 (2.1)</td>
<td>2 (0.5)</td>
<td>7 (1.6)</td>
<td></td>
</tr>
<tr>
<td><strong>Primary contact (n=341), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family member</td>
<td>309 (90.6)</td>
<td>63 (18.5)</td>
<td>246 (72.1)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>16 (4.7)</td>
<td>5 (1.5)</td>
<td>11 (3.2)</td>
<td></td>
</tr>
<tr>
<td>Health and social care professionals</td>
<td>15 (4.4)</td>
<td>1 (0.3)</td>
<td>14 (4.1)</td>
<td></td>
</tr>
<tr>
<td>Staff of living facility</td>
<td>1 (0.3)</td>
<td>0 (0)</td>
<td>1 (0.3)</td>
<td></td>
</tr>
</tbody>
</table>

*aNot applicable.

*bOne-way ANOVA.

*cChi-square test.

*dFor example, African, Haitian, Jamaican, or Somali.

eArab or West Asian (eg, Armenian, Egyptian, or Iranian), Latin American, South Asian, Korean, Mediterranean, Aboriginal (eg, Inuit, Métis, or American Indian), Filipino, Caribbean or West Indian (St Lucian or Antiguan), Southeast Asian, and Japanese.

### Table 2

Table 2 shows the history of missing incidents among MedicAlert subscribers. Subscribers self-reported no prior history of missing incidents at the time of subscribing to MedicAlert in 10.4% (45/433) of the cases, while 89.6% (388/433) disclosed having been involved in missing incidents repeatedly. Surprisingly, data from actual repeated missing incidents (ie, data that we accessed using the hotline access database) showed the opposite pattern: most of the subscribers went missing repeatedly in only 16.4% (71/434) of the cases. MedicAlert subscribers self-reported that the most common places to be found were outdoor public spaces (eg, highway or street; 184/308, 59.7%) or indoor public or communal spaces (eg, libraries; 85/308, 27.6%).
Table 2. Missing incidents history (unit of analysis: MedicAlert subscriber).

<table>
<thead>
<tr>
<th>Missing incidents history</th>
<th>Persons without dementia and persons living with dementia, n (%)</th>
<th>Persons without dementia, n (%)</th>
<th>Persons living with dementia, n (%)</th>
<th>Statistical test (persons without dementia vs persons living with dementia), chi-square (df)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing incident history (self-reported; n=433)</td>
<td>0.9 (2)</td>
<td>.64</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>45 (10.4)</td>
<td>7 (1.6)</td>
<td>38 (8.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeated (1^a-4 times)</td>
<td>327 (75.5)</td>
<td>67 (15.5)</td>
<td>260 (60)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Habitual (&gt;4 times)</td>
<td>61 (14.1)</td>
<td>14 (3.2)</td>
<td>47 (10.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Repeated missing incident history (actual; n=434)</td>
<td>2.1 (1)</td>
<td>.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>363 (83.6)</td>
<td>79 (18.2)</td>
<td>284 (65.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>71 (16.4)</td>
<td>10 (2.3)</td>
<td>61 (14.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possible locations for the missing person to be found (self-reported; n=308)</td>
<td>4.4 (3)</td>
<td>.22</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outdoor public space^b</td>
<td>184 (59.7)</td>
<td>34 (11)</td>
<td>150 (48.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indoor public or communal space^c</td>
<td>85 (27.6)</td>
<td>18 (5.8)</td>
<td>67 (21.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private home in the community^d</td>
<td>32 (10.4)</td>
<td>11 (3.8)</td>
<td>21 (6.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital (day program or day clinic)</td>
<td>7 (2.3)</td>
<td>1 (0.3)</td>
<td>6 (1.9)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^aAfter the first incident.

^bHighway, street, alley, intersection, park, parking lot, outdoor bus stop, construction, or cemetery.

^cGrocery store, shopping mall, train station, church, recreation center, library, physician’s office, bus, train, police station, or gas station.

^dHouse, apartment, condominium, age ≥65 years condominium but without supportive living services.

The statistical tests in Tables 1 and 2 showed that MedicAlert subscribers with dementia and those without who were involved in missing incidents were similar with respect to mean age, ethnic background, official Canadian languages spoken, province of residence, population density, living arrangement, primary contact, possible location to be found during a missing incident (self-reported), missing incident history (self-reported), and repeated missing incident history (actual missing incidents); no statistically significant differences were found between the groups. MedicAlert subscribers with dementia and those without who were involved in missing incidents are significantly different (χ² = 7.8; P = .05) in terms of age groups. This result means that the prevalence of critical wandering was higher among older age groups of people with dementia, with the peak prevalence between ages 75 and 84 years declining somewhat among the older-aged adults.

Characteristics of the Missing Incidents

**Demographic and Psychopathological Antecedents**

Missing incidents mostly involved people living with dementia (345/434, 79.6%) and those who were (1) female (230/431, 53.4%); (2) from the age groups 65-74 years (72/434, 16.6%), 75-84 years (177/434, 40.8%), and 85-94 years (168/434, 38.7%); (3) White (329/425, 77.4%); (4) English speaking (345/430, 80.2%); (5) living in an urban area (331/341, 97.1%); (6) mostly from Ontario (199/341, 58.4%), British Columbia (57/341, 16.7%), and Quebec (50/341, 14.7%); and (7) living with a family member (277/433, 64%) or in a facility (90/433, 20.8%).

**Locations**

Locations were analyzed in terms of point last seen or where the MedicAlert subscribers were missing from and located. MedicAlert subscribers were most frequently missing from private homes in the community (96/143, 67.1%) or residential living facilities (27/143, 18.9%); there were no statistical differences regarding the locations from which MedicAlert subscribers with dementia and those without went missing. Regarding the locations they were found, the most common places (313/382, 82%) were outdoor and indoor public spaces. Importantly, we found a statistically significant difference between point last seen or where the MedicAlert subscribers were missing from and where they were located (χ² = 42.3; P = .02; refer to Table 3 for more details). This result indicates that the MedicAlert subscribers were relatively active, with some degree of mobility. Even more interestingly, we found a moderate positive association between the possible locations to be found (self-reported) and the actual location where the person was found (Cramer V=0.213; P=.002).
Table 3. Characteristics of missing incidents. Point last seen or where the person was missing from and location in which the person was found (unit of analysis: missing incidents).

<table>
<thead>
<tr>
<th>Characteristics of missing incidents (locations)</th>
<th>Persons without dementia and persons living with dementia, n (%)</th>
<th>Persons without dementia, n (%)</th>
<th>Persons living with dementia, n (%)</th>
<th>Statistical test (persons without dementia vs persons living with dementia), chi-square (df)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Point last seen or where the person was missing from (n=143)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.64</td>
</tr>
<tr>
<td>Private home in the community</td>
<td>96 (67.1)</td>
<td>20 (14.8)</td>
<td>76 (51.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residential living facility(^a)</td>
<td>27 (18.9)</td>
<td>5 (4.0)</td>
<td>22 (15.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indoor public space(^b)</td>
<td>8 (5.6)</td>
<td>1 (0.7)</td>
<td>7 (4.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>5 (3.5)</td>
<td>2 (1.3)</td>
<td>3 (2.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital (day program or day clinic)</td>
<td>5 (3.5)</td>
<td>0 (0.0)</td>
<td>5 (3.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outdoor public space(^c)</td>
<td>2 (1.4)</td>
<td>0 (0.0)</td>
<td>2 (1.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Location in which the person was found (n=382)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.12</td>
</tr>
<tr>
<td>Outdoor public space</td>
<td>202 (52.9)</td>
<td>42 (11.3)</td>
<td>160 (40.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indoor public space</td>
<td>111 (29.1)</td>
<td>27 (7.8)</td>
<td>84 (22.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private home in the community</td>
<td>40 (10.5)</td>
<td>5 (1.3)</td>
<td>35 (8.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital (day program or day clinic)</td>
<td>12 (3.1)</td>
<td>2 (1.0)</td>
<td>10 (2.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residential living facility</td>
<td>11 (2.9)</td>
<td>2 (0.5)</td>
<td>9 (2.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>4 (1.0)</td>
<td>3 (0.8)</td>
<td>1 (0.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combination of &gt;1 of the aforementioned locations</td>
<td>2 (0.5)</td>
<td>0 (0.0)</td>
<td>2 (0.5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Long-term care center, assisted living facility, supportive living facility, lodge, or group home.
\(^b\)Grocery store, shopping mall, train station, church, recreation center, library, or physician’s office.
\(^c\)For example, highway, street, alley, intersection, park, or parking lot.

**Mode of Mobility**

While missing, the most common mode of traveling was on foot (248/270, 91.9%). The second most common mode of mobility was public transportation (12/270, 4.4%; refer to Table 4 for more details).
Table 4. Characteristics of missing incidents. Mode of mobility (unit of analysis: missing incidents; valid cases n=270).

<table>
<thead>
<tr>
<th>Mode of mobility</th>
<th>Persons without dementia and persons living with dementia, n (%)</th>
<th>Persons without dementia, n (%)</th>
<th>Persons living with dementia, n (%)</th>
<th>Statistical test (persons without dementia vs persons living with dementia), chi-square (df)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>On foot(^a)</td>
<td>248 (91.9)</td>
<td>57 (21.1)</td>
<td>191 (70.7)</td>
<td>2.3 (6)</td>
<td>.89</td>
</tr>
<tr>
<td>Public transit(^b)</td>
<td>12 (4.4)</td>
<td>3 (1.1)</td>
<td>9 (3.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receiving a ride from someone else(^c)</td>
<td>3 (1.1)</td>
<td>1 (0.4)</td>
<td>2 (0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driving own car</td>
<td>2 (0.7)</td>
<td>0 (0)</td>
<td>2 (0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>2 (0.7)</td>
<td>0 (0)</td>
<td>2 (0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combination of several of the aforementioned modes</td>
<td>2 (0.7)</td>
<td>0 (0)</td>
<td>2 (0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Long-range or transregional transit(^d)</td>
<td>1 (0.4)</td>
<td>0 (0)</td>
<td>1 (0.4)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Walking. 
\(^b\)Bus, light rail transit, subway, or streetcar. 
\(^c\)Hitchhiking. 
\(^d\)Train, airplane, noncommuter bus, or ferry.

Temporality (Time and Seasonality)

Table 5 shows the temporality of missing incidents in terms of the time of day and season the person was reported missing. In general, missing incidents occurred mostly in the afternoon (262/560, 46.8%) and the evening (174/560, 31.1%), while the most common seasons for these incidents were summer and fall followed by spring (222/560, 39.6%; 154/560, 27.5%; and 113/560, 20.2%, respectively). No statistical differences for the temporality variable by persons living with dementia and those without were found.

Table 5. Characteristics of missing incidents. Time and seasonality (unit of analysis: missing incidents; n=560).

<table>
<thead>
<tr>
<th>Missing incident characteristics (time and seasonality)</th>
<th>Persons without dementia and persons living with dementia, n (%)</th>
<th>Persons without dementia, n (%)</th>
<th>Persons living with dementia, n (%)</th>
<th>Statistical test (persons without dementia vs persons living with dementia), chi-square (df)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time of day</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afternoon (noon to 5:59 PM)</td>
<td>262 (46.8)</td>
<td>61 (11)</td>
<td>201 (35.1)</td>
<td>7.2 (2)</td>
<td>.03</td>
</tr>
<tr>
<td>Evening (6 PM to 11:59 PM)</td>
<td>174 (31.1)</td>
<td>40 (7.1)</td>
<td>134 (23.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morning (midnight to 11:59 AM)</td>
<td>124 (22.1)</td>
<td>15 (3.9)</td>
<td>109 (19.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Season</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Summer (June 1 to August 31)</td>
<td>222 (39.6)</td>
<td>47 (8.4)</td>
<td>175 (31.3)</td>
<td>3.3 (3)</td>
<td>.34</td>
</tr>
<tr>
<td>Fall (September 1 to November 30)</td>
<td>154 (27.5)</td>
<td>34 (6.1)</td>
<td>120 (21.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring (March 1 to May 31)</td>
<td>113 (20.2)</td>
<td>17 (3)</td>
<td>96 (17.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Winter (December 1 to February 28)</td>
<td>71 (12.7)</td>
<td>18 (3.2)</td>
<td>53 (9.5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

People Involved in the Missing Incident

Table 6 shows the people involved in the missing incidents in terms of the care partner involvement with MedicAlert in response to the missing incident. In the majority of cases (375/518, 72.4%), the family care partner had an involvement in response to the missing incident with MedicAlert, with no statistically significant difference found between the groups (people with dementia and those without). In 96.1% (467/486) of the cases, MedicAlert subscribers who went missing were located by someone other than the care partner. In most of the cases, either first responders (232/486, 47.7%) or Good
Samaritans (224/486, 46.1%) found the missing person. Again, no statistically significant difference was found between the groups involved in the missing incidents.

### Table 6. Characteristics of missing incidents. People involved in the missing incident (unit of analysis: missing incidents).

<table>
<thead>
<tr>
<th>Missing incident characteristics (people involved in the missing incident)</th>
<th>Persons without dementia and persons living with dementia, n (%)</th>
<th>Persons without dementia, n (%)</th>
<th>Persons living with dementia, n (%)</th>
<th>Statistical test (persons without dementia vs persons living with dementia), chi-square (df)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Natural care partner involvement in response to incident with MedicAlert (n=518)</td>
<td></td>
<td></td>
<td></td>
<td>0.4 (1)</td>
<td>.54</td>
</tr>
<tr>
<td>Yes</td>
<td>375 (72.4)</td>
<td>80 (15.4)</td>
<td>295 (57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>143 (27.6)</td>
<td>27 (5.2)</td>
<td>116 (22.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Who reported and found the missing person (n=486)</td>
<td></td>
<td></td>
<td></td>
<td>1.3 (3)</td>
<td>.72</td>
</tr>
<tr>
<td>First responder&lt;sup&gt;a&lt;/sup&gt;</td>
<td>232 (47.7)</td>
<td>54 (11.1)</td>
<td>178 (36.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good Samaritan&lt;sup&gt;b&lt;/sup&gt;</td>
<td>224 (46.1)</td>
<td>48 (9.9)</td>
<td>176 (36.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family member or friend&lt;sup&gt;c&lt;/sup&gt;</td>
<td>19 (3.9)</td>
<td>4 (0.8)</td>
<td>15 (3.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>11 (2.3)</td>
<td>1 (0.2)</td>
<td>10 (2.1)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>Police, search and rescue member, fire department, or ambulance or paramedic.

<sup>b</sup>The Good Samaritan noticed that something was off with the missing person and called the hotline or was asked by the missing person to call the hotline; they were not formally involved in searching for the missing person.

<sup>c</sup>Informal care partner.

### Outcomes of the Missing Incidents

Table 7 shows the outcomes of the missing incidents in terms of the number of missing incidents, repeated missing incidents, mean time to the first missing incident (in days), mean time between missing incidents (in days), and survivability. Overall, 22.5% (113/500) of the missing incidents were repeated missing incidents, with the mean number of missing incidents per MedicAlert subscriber being 1.290 (SD 0.914; range: 1-11). Moreover, the number of missing incidents per MedicAlert subscriber was almost the same for people living with dementia (mean 1.290, SD 0.801) and those without dementia (mean 1.300, SD 1.265). The mean time to the first missing incident (since registering with MedicAlert) was 343.8 (SD 376.2) days (mean 11, SD 11.3 months), whereas the mean time between missing incidents was shorter, that is, 328.0 (SD 366.6) days (mean 11, SD 10.8 months). This is expected because the mean time between missing incidents takes into account repeated missing incidents. In terms of survivability, only a small percentage of cases (46/500, 9%) involved people undergoing harm while missing; even more rare were missing incidents in which MedicAlert subscribers were found deceased (1/500, 0.2%). There was a trend toward adverse outcomes for MedicAlert subscribers living with dementia: they experienced increased repeated missing incidents and injuries (but these results were not statistically significant, \( P = .30 \)), short mean time to the first missing incident (but these results were not statistically significant, \( P = .20 \)), and short mean time between missing incidents (but these results were not statistically significant, \( P = .15 \)). In other words, they went missing more frequently (1 missing incident every 317.08 days) than those subscribers who did not have dementia (1 missing incident every 370.41).
Table 7. Outcomes of the missing incidents.

<table>
<thead>
<tr>
<th>Missing incident characteristics (outcomes)</th>
<th>Persons without dementia and persons living with dementia</th>
<th>Persons without dementia, n (%)</th>
<th>Persons living with dementia, n (%)</th>
<th>Statistical tests (persons without dementia vs persons living with dementia)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Missing incidents (n=434)</td>
<td>Mean 1.290 (SD 0.914; range 1-11)</td>
<td>1.300 (1.265)</td>
<td>1.290 (0.801)</td>
<td>Mann-Whitney U test</td>
<td>-1.386</td>
</tr>
<tr>
<td>Time (d; n=434), mean (SD)</td>
<td>Mean 343.79 (SD 376.20; range 6-2249)</td>
<td>374.82 (410.35)</td>
<td>335.79 (365.09)</td>
<td>z score</td>
<td></td>
</tr>
<tr>
<td>MTFIc (n=434)</td>
<td>Mean 328.02 (SD 366.62; range 6-2249)</td>
<td>370.41 (411.34)</td>
<td>317.08 (354.00)</td>
<td>Mann-Whitney U test</td>
<td>1.429</td>
</tr>
<tr>
<td>MTBI d (n=434)</td>
<td>Mean 343.79 (SD 376.20; range 6-2249)</td>
<td>374.82 (410.35)</td>
<td>335.79 (365.09)</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Survivability (n=500), n (%)

<table>
<thead>
<tr>
<th>Event</th>
<th>Persons without dementia</th>
<th>Persons living with dementia</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No apparent injuries or compromised health</td>
<td>453 (90.6)</td>
<td>94 (18.8)</td>
<td>359 (71.8)</td>
</tr>
<tr>
<td>Injuries or compromised health requiring emergency services and transfer to hospital</td>
<td>35 (7)</td>
<td>8 (1.6)</td>
<td>27 (5.4)</td>
</tr>
<tr>
<td>Minor injuries or health issues requiring some treatment at homee</td>
<td>10 (2)</td>
<td>1 (0.2)</td>
<td>9 (1.8)</td>
</tr>
<tr>
<td>Deceased</td>
<td>1 (0.2)</td>
<td>1 (0.2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Injuries and concern for health requiring follow-up caref</td>
<td>1 (0.2)</td>
<td>0 (0)</td>
<td>1 (0.2)</td>
</tr>
</tbody>
</table>

aNot applicable.
bMann-Whitney U test.
cMTFI: mean time to the first missing incident.
dMTBI: mean time between missing incidents.
eChi-square test.
fGetting Band-Aids, pain medications, cleaned up, and so on.

Discussion

Principal Findings

This retrospective descriptive study examined demographic, psychopathological, and environmental antecedents to missing incidents due to critical wandering among MedicAlert subscribers, as well as the characteristics and outcomes of these incidents. In doing so, we used a national registry of persons as a secondary data source of information (ie, the MedicAlert database). To date, much of the knowledge about missing individuals with dementia and those without is based on studies with small sample sizes that use social media and newspaper reports from the United States or elsewhere [7,11]. Thus, we aimed to address these limitations by using an extensive secondary data set. To our knowledge, this is the first study that has shed light on the phenomenon of missingness and critical wandering of individuals with dementia and those without in Canada. In addition, we were able to report the prevalence of repeated missing incidents, based on this database, an important figure that has been absent in previous studies.

The demographic characteristics of our study population showed that people involved in missing incidents were mostly older adults (345/434, 79.6%), female older adults (230/431, 53.4%), living in the most populated provinces in Canada (306/341, 89.7%), and living in urban areas with at least 1 family member (309/341, 90.6%). Importantly, the majority of MedicAlert subscribers (345/434, 79.5%) involved in missing incidents self-reported living with dementia. More interestingly, except for age group, we did not find statistically significant differences between the people living with dementia and those without with
The literature reports that neurocognitive deficits could include memory deficits, such as declarative memory (remembering facts and events), episodic memory (short-term memory for recent events and contexts), and visual agnosia (inability to recognize objects or places). In addition, executive function impairments and disease-related changes to visuospatial and subperceptual processing (especially in unfamiliar locations), which typically manifests as difficulty with navigation, can explain why an individual living with dementia cannot independently return home [26].

The prevalence of MedicAlert subscribers who repeatedly went missing was lower in the hotline database in comparison to repeated missing self-reported variables. This result was anticipated because previous studies suggest that care partners are reluctant to contact emergency services, such as 911 or programs to locate older adults who are having an episode of critical wandering and have gone missing [37]. As a self-reported variable, this result could be attributed to overreporting by care partners. However, more objective explanations can be given. First, care partners often initiate the search within their homes or places last seen, and because many persons with dementia are found near the place that they were last seen, either on their own property or in their own neighborhood, the care partner could locate the missing person before their decision to request assistance from external organizations [31,37]. Second, it is possible that care partners subscribed their family member into the MedicAlert program as a preventative measure. For individuals who repeatedly had episodes of critical wandering in the past, care partners could have implemented their own measures or interventions to avoid missing incidents, including MedicAlert subscriptions. The literature reports that these interventions include avoiding lapses of supervision, whether planned or unplanned, through the use of technology (eg, GPS) to monitor and locate missing older adults with dementia [39]. Finally, it might be possible that care partners chose not to use the MedicAlert hotline to locate missing individuals to avoid attention and stigma associated with a formal search if initiated. Numerous studies have reported that the uses of technologies and programs by people living with dementia and their care partners aiming to reduce the risks of getting lost have highlighted the importance of discreet technologies that are unnoticeable to reduce stigma [40,41].

Our study paves the way for new services and interventions that can be offered by MedicAlert. The services may include implementing preventative strategies to decrease the risk of going missing through threshold alerts in mobile phone apps. According to the literature, a leading feature being implemented and the mean time to the first missing incident for a particular MedicAlert subscribers is approaching. The same can be true for common months or the time of day that MedicAlert subscribers are reluctant to contact emergency services, such as 911 or programs to locate older adults who are having an episode of critical wandering and have gone missing [37].

We found that the majority of MedicAlert subscribers involved in missing incidents self-reported living with dementia (345/434, 79.5%). Importantly, among those who were involved in missing incidents but did not self-report living with dementia (89/434, 20.1%), memory loss was self-reported as a medical condition. Our result is aligned with previous studies that found that persons with mental or cognitive disabilities (eg, those with Alzheimer disease or dementia) are more prone to going missing [31,37,38]. The literature reports that neurocognitive deficits from dementia predispose individuals to missing incidents and contribute to the inability to independently return home. These neurocognitive deficits could include memory deficits, such as declarative memory (remembering facts and events), episodic memory (short-term memory for recent events and contexts), and visual agnosia (inability to recognize objects or places). In addition, executive function impairments and disease-related changes to visuospatial and subperceptual processing (especially in unfamiliar locations), which typically manifests as difficulty with navigation, can explain why an individual living with dementia cannot independently return home [26].
engages community citizens as volunteers to look for missing persons with dementia, has demonstrated to be an accurate and useful app [43].

For MedicAlert subscribers involved in missing incidents, many were returned home safely (453/500, 90.6%), with few reported harms or injuries (46/500, 9.2%), and death was rare (1/500, 0.2%). Regarding mortality rates when a person with dementia goes missing, the literature shows high variability (between 0.7% and 32%) [44,45]. In this study, the low reports of harm and death can be explained mainly by 2 factors. First, the environmental conditions at the time the MedicAlert subscribers went missing were favorable: subscribers went missing in urban areas while traveling on foot (248/270, 91.9%) or using public transportation (12/270, 4.4%) during the day (262/560, 46.8%) in the warmest months of the year (eg, low chance of severe weather; 358/560, 63.9%). The literature reports that the causes of high mortality rates in people with dementia who go missing include severe weather; driving; and walking near roadways, bodies of water, or in isolated areas [46,47]. These scenarios were very different from what we found in our study. Second, we found that in a high proportion of missing incidents (504/560, 90%), the MedicAlert subscribers were wearing their ID bracelet. We could intuitively affirm that the MedicAlert program prevents injuries and saves lives, but this affirmation would have to be demonstrated in a formal study. Therefore, a next logical step would be to conduct a study to determine whether the MedicAlert program addresses the problem for which it was designed, that is, to help those who are having an episode of critical wandering to return home safely. Our study also shows what some investigators have determined regarding the potential interrelatedness of risk factors for going missing [48]. While most of the outcomes during missing incidents were positive (death was rare), the complex interplay of demographic, psychopathological, and environmental antecedents of MedicAlert subscribers need further exploration.

Study Limitations

Our study has some limitations. Limitations were posed by the MedicAlert data set itself. First, missing incidents are also captured in data held by first responders (police, search and rescue organizations, paramedics, etc), and, because MedicAlert data are subscription-based data, there are inherent self-selection biases. Second, while inquiring about the data entry process at the MedicAlert subscribers’ level, we discovered that a high percentage of data were stored raw (free-text fields) and not in analysis-ready format. Consequently, the available information did not allow us to categorize our data with the desired level of granularity. Third, the self-reported nature of the data caused missing data in some variables (eg, the use of de-escalation techniques to avoid missing incidents, whether a MedicAlert subscriber has special needs, and what constitutes a trigger for a missing incident). As the percentage of missing data in these variables was large (ie, >40%), we excluded them from the analyses as recommended in the literature [49]. The missing data will not allow for further comprehensive statistical analysis for these unmeasured confounding variables. Fourth, the database lacked some important outcome variables; for example, we were unable to determine for how long MedicAlert subscribers went missing, the response time (ie, time elapsed between the call to the hotline and the arrival of the first responders to assist a missing person) or the turnaround time of the missing incidents (ie, the time it takes to return a missing person to their residence). In summary, because this study used a secondary data source that had not been compiled for research purposes, we faced the same common limitations reported in other studies that use this kind of data source [50-52]. Notwithstanding these limitations, we believe that the results obtained in this study are very valuable for partially understanding the phenomenon of older adults with dementia and memory loss going missing in Canada. The data set used in this study represents a small portion of people living with dementia in Canada; by virtue of it being a paid subscription service, not everyone uses it. In future research, other sources of data also need to be considered (police and search and rescue data) to get a fuller picture of the prevalence of persons living with dementia who go missing.

Conclusions

In the data set used, missing incidents involved mostly female older adults living with dementia from an urban area (331/341, 97.1%). Overall, the majority of MedicAlert subscribers involved in missing incidents returned home safely (453/500, 90.6%). However, 9.2% (46/500) of the cases resulted in some form of minor or serious injuries and death. Of the 560 missing incidents, 126 (22.5%) were repeated missing incidents. This paves the way to more accurately describe the prevalence of missing incidents and their consequences and outcomes so that we can develop targeted intervention strategies to prevent missing incidents or locate missing persons.

Acknowledgments

This work was supported by the Search and Rescue New Initiatives Fund program, Public Safety Canada (22005). The authors thank Serrina Philip, Marria Khalid, Andrew Faller-Saunders, and Vanessa Vahedi for their help in data cleaning and the categorization process. The authors also thank Elea Thullier for her help in conducting some of the data analyses. Finally, the authors thank the MedicAlert team for their help and support in understanding the MedicAlert database.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Definitions of variables and measures.
References


The Prevalence of Missing Incidents and Their Antecedents Among Older Adult MedicAlert Subscribers: Retrospective Descriptive Study

Miguel-Cruz et al. (2024) published an article titled "The Prevalence of Missing Incidents and Their Antecedents Among Older Adult MedicAlert Subscribers: Retrospective Descriptive Study" in JMIR Aging. The study investigated the prevalence of missing incidents among older adult MedicAlert subscribers and identified their antecedents. The findings are relevant for understanding the safety concerns and preventive measures for older adults using MedicAlert services. This work contributes to the field of geriatric medicine by highlighting the importance of monitoring and addressing missing incidents among this population.

©Antonio Miguel-Cruz, Hector Perez, Yoojin Choi, Emily Rutledge, Christine Daum, Lili Liu. Originally published in JMIR Aging (https://aging.jmir.org), 10.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Refining Cultural Adaptations of a Behavioral Intervention for Latino Caregivers of People Living With Dementia: Qualitative Interview Study in Washington State

Celeste N Garcia¹, BA; Miriana C Duran¹, MPH, MD; Magaly Ramirez¹, MS, PhD
Department of Health Systems and Population Health, University of Washington, Seattle, WA, United States

Corresponding Author:
Magaly Ramirez, MS, PhD
Department of Health Systems and Population Health
University of Washington
3980 15th Avenue North East
Seattle, WA, 98195
United States
Phone: 1 2065439773
Email: maggiera@uw.edu

Abstract

Background: In the United States, Latino caregivers of individuals with dementia face unique challenges and an elevated risk of adverse health outcomes. Despite the increasing prevalence of Alzheimer disease and related dementias among Latino adults, few evidence-based interventions are tailored to their cultural context. To address this gap, we examined the cultural adaptations required for the STAR caregivers (STAR-C) virtual intervention, an evidence-based intervention that educates family caregivers to manage behavioral and psychological symptoms of dementia. While STAR-C has shown effectiveness, neither the original in-person nor the virtual intervention considered the distinct experiences of Latino caregivers, who often bring culturally significant values into caregiving interactions.

Objective: This study’s objective was to test and refine the preliminary cultural adaptations of the STAR-C web-based training modules for Latino caregivers of people living with dementia.

Methods: Through qualitative interviews with 15 Latino caregivers in Washington State, we identified key adaptations to enhance the cultural relevance of the web-based training modules.

Results: The interviews highlighted 4 main themes for adaptation: the delivery of the STAR-C web-based training modules, comprehensive dementia education, simplified problem-solving strategies, and prioritizing caregiver well-being.

Conclusions: This study’s findings informed the development of culturally adapted STAR-C web-based training modules that aim to provide tailored support to Latino caregivers. While further research is needed to assess the efficacy of these adaptations, our work contributes to bridging the gap in dementia caregiving for Latino families, potentially reducing health disparities and enhancing health care services for this population.

Introduction

Latino caregivers of people living with dementia are at an increased risk of experiencing adverse health impacts due to caregiving, yet few evidence-based interventions have been developed to support Latino families [1]. In the United States, Latino adults are 1.5 times more likely to develop Alzheimer disease and related dementias (ADRD) compared to non-Latino
White adults [2]. The disparity is due in large part to the health conditions (eg, cardiovascular disease, diabetes, high blood pressure, and obesity) and socioeconomic factors (eg, chronic exposure to economic and social adversity, lower levels and quality of education, and discrimination) that are more prevalent in Latino populations and are associated with cognitive decline [2,3]. The number of Latino people living with dementia is expected to increase to 3.5 million by 2060, leading to a rise in Latino adults caring for family members with ADRD [4]. Although evidence-based caregiver interventions exist, they often fall short in meeting the unique cultural needs of Latino families [1,5,6]. There is an urgent need to develop culturally appropriate evidence-based interventions that address the unique challenges faced by Latino caregivers of people living with dementia and consider the sociocultural context in which they provide care.

STAR caregivers (STAR-C) is an in-home intervention that involves training health professionals to teach family caregivers strategies to manage behavioral and psychological symptoms of dementia (BPSD) [7,8]. Caregivers learn to monitor symptoms, identify possible environmental or interpersonal triggers, and develop effective responses. They also learn strategies for communicating with people living with dementia in a way that supports positive affect and prevents or minimizes problems, increasing pleasant events to improve mood, and improving the support caregivers receive from informal and formal networks. STAR-C is demonstrated to reduce the frequency and severity of BPSD, as well as improve burden, depression, and reactivity to symptoms in caregivers [7]. Recently, STAR-C was reconfigured as a virtual intervention to facilitate large-scale implementation in clinical settings. The virtual intervention, coined STAR-C Virtual Training and Follow-up, is being tested in an ongoing trial at Kaiser Permanente Washington [9,10]. For 6-8 weeks, caregivers complete 6 web-based training modules asynchronously and have six 30-minute weekly telephone check-ins with a coach (ie, master’s-level social worker or mental health counselor). In addition, support from coaches is provided, as needed, via secure messaging in the Kaiser Permanente Washington patient portal for up to 6 months.

The STAR-C virtual intervention was timely given the COVID-19 pandemic, which shed light on the urgent need for digital health strategies that offer support virtually [11]. Many social and health care services for older adults and their family caregivers shifted from in-person to digital platforms to expand reach during the pandemic [12]. Neither the in-person or virtual STAR-C interventions, however, were developed with explicit consideration of the experience of Latinos providing care to a family member living with ADRD. Caregiver interventions for Latinos need cultural adaptations because cultural values and family) and respeto (respect) play a pivotal role in shaping caregiving interactions with people living with dementia, experiences, and perceptions of support [13]. The goal of performing cultural adaptations to evidence-based interventions such as STAR-C is to promote more favorable experiences with the intervention and alleviate the health disparities associated with dementia caregiving among vulnerable populations.

To address the gap in the lack of culturally appropriate evidence-based interventions for Latino caregivers, we sought to culturally adapt the content of the web-based training modules of the STAR-C virtual intervention for Latino caregivers. In our previous study, we identified what cultural adaptations to the STAR-C web-based training modules are needed for Latino caregivers and we designed preliminary adaptations [14]. Preliminary adaptations included expanding the content of the web-based training modules to improve understanding of dementia; revising language that was viewed as stigmatizing, offensive, or culturally inappropriate; and adding cultural examples to reflect the range of family involvement in caring for people living with dementia and multigenerational living [14]. It is unknown, however, whether these modifications sufficiently align with the intended goals of cultural adaptation [15]. In addition, it is unknown whether there are additional opportunities to further enhance the cultural relevance of the STAR-C web-based training modules for Latino families. Therefore, the objective of this study was to test and refine the preliminary cultural adaptations of the STAR-C web-based training modules for Latino caregivers of people living with dementia.

**Methods**

**Ethical Considerations**

This study was granted approval by the institutional review board at the University of Washington (STUDY00009534). Participants in this study gave their verbal or written consent for their involvement. All the data we collected from participants was labeled with a unique study identification number and not the participants’ name or any other information that could identify participants. The contact information of participants was kept in a password-protected file and computer. All data collected from participants were kept confidential and accessible only by our study team. We did not use participants’ names in reports of study findings, REDCap (Research Electronic Data Capture; Vanderbilt University) surveys, or audio recordings of interviews. Instead, we labeled everything with this study’s identification number. We destroyed data that identified participants when we finished recruitment.

**Overview of This Study’s Design**

Figure 1 [16] illustrates our approach to the cultural adaptation of the STAR-C web-based training modules within the context of the Discover, Design + Build, Test framework. In our previous study, we conducted a qualitative study to gather information about needed cultural adaptations to the web-based training modules for Latino caregivers (the “Discover” phase) [14]. We then used the findings from the qualitative study to brainstorm ideas for preliminary cultural adaptations (the “Design” phase). In this study, we developed low-fidelity prototypes of the culturally adapted STAR-C web-based training modules and tested the prototypes with Latino caregivers (the “Build” phase).
Participant Selection

We carried out semistructured qualitative interviews in Washington State with 15 Latino caregivers of people living with dementia who spoke Spanish or English. Caregivers were considered eligible if they self-identified as Hispanic or Latino, were aged 21 years or older, were a child, spouse, or partner, or close friend of someone diagnosed with dementia, lived with the diagnosed individual or within a 5-mile radius, and provided a minimum of 8 hours of weekly care. Our recruitment strategy involved 3 approaches. Initially, we identified potential participants through the electronic health record system at University of Washington Medicine. Additionally, we circulated flyers in both Spanish and English across various locations, including a UW Medicine specialty clinic, a primary care practice-based research network, the Alzheimer Association Washington State Chapter, and local tiendas in Latino communities. Lastly, we used media platforms like local Spanish and English radio stations and newspapers to discuss the impact of dementia on the Latino population and publicize this study. A member of this study’s team screened for eligibility the individuals who showed interest in joining this study and arranged interviews for those who met the eligibility requirements. Caregivers received a compensation of US $45 for taking part in this study.

Description of the Low-Fidelity Prototypes

In the STAR-C Virtual Training and Follow-up intervention, caregivers receive 1 web-based training module per week for a period of 6-8 weeks. Textbox 1 describes the topics of the web-based training modules. The core components of the intervention include dementia education, strategies for effective communication, Activators-Behaviors-Consequences (ABC) problem-solving, pleasant events, and caregiver support.

Textbox 1. Topics of STAR caregivers (STAR-C) web-based training module.

| Week 1: Understanding dementia, realistic expectations about behavioral treatments for behavioral and psychological symptoms of dementia, and strategies for effective communication. |
| Week 2: Activator, Behavior, Consequence (ABC) approach to problem-solving, including rationale and development of an ABC plan for target behaviors that caregivers identify. |
| Week 3: Review of ABC plan (revise if needed). |
| Week 4: Pleasant events and managing negative thinking. |
| Week 5: Review of ABC plan and pleasant events schedule (revise if needed). |
| Week 6: Caregiver support strategies for coping with caregiving and maintaining gains. |

We created 3 low-fidelity prototypes of the culturally adapted STAR-C web-based training modules in both English and Spanish for testing among study participants. The low-fidelity prototypes were in the form of videos of recorded presentations with images, text, and voice-over. The videos featured short excerpts from various STAR-C web-based training modules. The first video consisted of educating Latino family caregivers on dementia and problem-solving using the ABC approach. This video explained the STAR-C program, dementia, its causes, walked through each step of the ABC approach for problem-solving in dementia, and concluded with a caregiving example of the ABC approach. The second video featured some modifications, including the addition of information about dementia stages, common behaviors exhibited by people living with dementia, and importance of self-care. We also enhanced the visual aesthetics and design of the lessons to make them more visually appealing and added humor. In the third video, we maintained the core content from the previous versions but incorporated interactive images and examples to increase engagement and interactivity. Figures 2 and 3 provide an example of the low-fidelity prototypes.
Figure 2. Culturally adapted STAR-C web-based training module—understanding Alzheimer and related dementias, with a focus on disproportionate impact on Latinos. STAR-C: STAR caregiver.

What is dementia?

- Dementia impacts many families in our community.
- Latinos are more likely that other racial or ethnic groups to get dementia.
- In the US, 1 out of 8 Latinos over 65 has dementia.

Figure 3. Culturally adapted STAR-C web-based training module—ABC acronym. ABC: Activators-Behaviors-Consequences; STAR-C: STAR caregiver.

ABC problem-solving

The ABCs of how to solve problems you face when caring for someone with memory loss.

The goals of this session are to:
- Introduce the ABCs:
  - Activators (causes)
  - Behaviors
  - Consequences
- Come up with specific plans to help change behaviors.

Data Collection

Study participants received the low-fidelity prototypes (ie, videos) via text message, WhatsApp Messenger, or email, depending on their preference, and were asked to watch the video at home on their own time prior to the scheduled qualitative interview. A bilingual or bicultural staff member trained in qualitative research techniques conducted interviews from January to August 2022. These interviews were held virtually, and they took place in either Spanish or English based on the participant’s choice. Each interview spanned 30 to 60 minutes in length and was audio recorded and transcribed verbatim by a professional service.

The staff member used semistructured interview guides. For the first 5 interviews, the interview guide asked general questions pertaining to likes and dislikes, queries about the program’s dementia content, examination of the ABC problem-solving strategy, assessment of the provided caregiving example, and a series of questions delving into video accessibility, design, and duration. For the subsequent 5 interviews, we revised the interview guide by adding questions focused on participant preferences, such as whether they preferred watching or listening to the video content, their favored method of receiving the videos, and their ideal video length. Finally, after another subsequent 5 interviews, we revised the interview guide again by incorporating additional inquiries centered around evaluating the images within the videos and improving viewer engagement.

Data Analysis

We analyzed the transcripts in their original languages, either Spanish or English, using the qualitative data analysis techniques described by Saldaña [17]. We began by reading through all the transcripts and writing analytic memos to reflect on the
content—noting insights, thoughts, and emerging patterns. We then applied deductive codes from a pre-established codebook to the interview transcripts (“protocol coding” [17]). The pre-established codebook was used in a prior qualitative study of STAR-C and was composed of codes representing components of the cultural treatment adaptation framework [14,18]. As an illustration, the codebook contained codes like “Materials and Semantics,” “Cultural Examples and Themes,” and “Therapeutic Framework” to capture cultural adaptations that were needed in the delivery of the intervention. In the next step, we applied inductive second-order codes to capture the details of participants’ feedback (“subcoding” [17]). For example, one of the inductive subcodes under “Materials and Semantics” was “Different BPSD examples” since multiple participants had suggested including in the intervention materials problem-solving examples for different BPSD. Afterward, we grouped the deductive and inductive codes into a smaller number of candidate themes that represented opportunities to improve the cultural adaptations of STAR-C (“pattern coding” [17]). We examined the interrelationship across and within the candidate themes and made refinements to ensure that excerpts within themes cohered and that each final theme was distinct from the others. Finally, we developed a statement to describe each of themes, which are presented in the following section. To manage this coding process, we used Dedoose (version 8.1.8; SocioCultural Research Consultants, LLC).

**Results**

**Overview**

Table 1 provides the characteristics of Latino caregivers who participated in this study and a description of their caregiving situation. The average age of caregivers was 48.9 (SD 11.1) years, 13 (87%) identified as women, 7 (47%) were the adult children of the people living with dementia, and 12 (80%) provided 35 hours or more of care per week.
Table 1. Characteristics of family caregivers and description of the caregiving situation.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Family caregivers (N=15), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>48.9 (11.1)</td>
</tr>
<tr>
<td>Mexican, Mexican American, or Chicano</td>
<td>15 (100)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Woman</td>
<td>13 (87)</td>
</tr>
<tr>
<td>Man</td>
<td>2 (13)</td>
</tr>
<tr>
<td><strong>Occupational status</strong></td>
<td></td>
</tr>
<tr>
<td>Retired</td>
<td>1 (7)</td>
</tr>
<tr>
<td>Employed</td>
<td>5 (33)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>1 (7)</td>
</tr>
<tr>
<td>Homemaker</td>
<td>5 (33)</td>
</tr>
<tr>
<td>Other</td>
<td>3 (20)</td>
</tr>
<tr>
<td><strong>Highest level of educational attainment, mean (SD)</strong></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>4 (27)</td>
</tr>
<tr>
<td>High school</td>
<td>5 (33)</td>
</tr>
<tr>
<td>Vocational or technical training</td>
<td>2 (13)</td>
</tr>
<tr>
<td>Some college</td>
<td>1 (7)</td>
</tr>
<tr>
<td>College graduate</td>
<td>2 (13)</td>
</tr>
<tr>
<td>Postgraduate</td>
<td>1 (7)</td>
</tr>
<tr>
<td><strong>Devices owned</strong></td>
<td></td>
</tr>
<tr>
<td>Smartphone</td>
<td>14 (93)</td>
</tr>
<tr>
<td>Tablet</td>
<td>5 (33)</td>
</tr>
<tr>
<td>Laptop</td>
<td>7 (47)</td>
</tr>
<tr>
<td>Computer</td>
<td>6 (40)</td>
</tr>
<tr>
<td>Owns any device</td>
<td>15 (100)</td>
</tr>
<tr>
<td><strong>Caregiver’s relationship to person living with dementia, mean (SD)</strong></td>
<td></td>
</tr>
<tr>
<td>Adult child (eg, daughter)</td>
<td>7 (46.7)</td>
</tr>
<tr>
<td>Spouse or partner</td>
<td>5 (33.3)</td>
</tr>
<tr>
<td>Other relative</td>
<td>3 (20)</td>
</tr>
<tr>
<td><strong>Care provided per week (hours)</strong></td>
<td></td>
</tr>
<tr>
<td>35 or more</td>
<td>12 (80)</td>
</tr>
<tr>
<td>5-14</td>
<td>3 (20)</td>
</tr>
<tr>
<td>Number of years providing care, mean (SD)</td>
<td>3.7 (2.4)</td>
</tr>
<tr>
<td>Caregiver and person living with dementia living together, mean (SD)</td>
<td>10 (66.7)</td>
</tr>
</tbody>
</table>

The qualitative analysis revealed adaptations that were needed in the (1) delivery of the STAR-C virtual intervention, (2) “dementia education” core component, (3) “ABC Problem-Solving” core component, and (4) “caregiver support strategies” core component. The sections below describe the need for these adaptations from the perspective of caregivers, as well as the changes that were made to the STAR-C web-based training modules in response to caregiver feedback.

**Theme 1: Adaptations to the Delivery of the STAR-C Virtual Intervention**

**Design Videos to be Accessible via Audio Only**

Latino caregivers reported that they liked that the STAR-C videos combined visuals with narration. They reported that the narration helped to reinforce the visual content and vice versa. They said this combination of visuals and narration was engaging because it helped to hold their attention and made it easier for them to learn the content. For example, 1 Spanish-speaking caregiver said:
The thing is that audio and seeing the image are also easier in this type of application. It became more practical and easier for me to learn and understand compared to just reading and seeing it without any—well, at least in my opinion, my brain doesn’t work the same way, I believe, but I liked the idea that you can see it, like imagine it, and create that idea.

However, while the ideal would be to watch the STAR-C videos, some caregivers mentioned that this may not always be practical. They said that they were often occupied with errands, household, and caregiving work, and that it would be nice in these circumstances to be able to listen to the video. For that reason, the caregivers suggested that the videos be designed to be accessible via audio only. Further, 1 Spanish-speaking caregiver stated:

Since sometimes one doesn’t have much time to sit down and take the time to be looking at the screen. So, I put everything in the background while I do my daily tasks.

In response to caregivers’ feedback, we modified the narration script to ensure that it was independent of visuals and accessible via audio only. The narration focused on providing clear and descriptive explanations of the visual content. Instead of relying on visuals to convey information, the narration script described the key elements, actions, and visuals present in the video. This approach would enable caregivers to form a mental image and grasp the content without needing to see the visuals.

Make Videos Accessible on Multiple Platforms

Latino caregivers expressed a preference for having various methods of accessing the STAR-C videos, including phone applications such as Facebook Messenger, WhatsApp Messenger, YouTube, and other platforms. Caregivers desired video access through platforms they used regularly and were easily accessible on their preferred devices. For example, 1 Spanish-speaking caregiver stated:

No, but I hardly use my email. I do check my email, but almost all the information I receive, I receive it through WhatsApp. And then, I can see it right away without any problem.

A few caregivers encountered technical difficulties when attempting to access the videos such as lack of Wi-Fi access in their homes or limited storage on their phones, requiring troubleshooting. For example, 1 caregiver said she could not view the videos initially because her phone’s storage was full. Once the caregiver emptied her phone’s storage, she was able to view the videos. In addition, while the caregiver could view the videos on her computer, she did not have Wi-Fi at home and would need to connect the computer to her phone’s hotspot.

In response to caregivers’ feedback, we considered various options to accommodate preferences and technology access, to ensure optimal accessibility to the STAR-C videos. We would offer caregivers the option to receive the STAR-C videos on phone applications like Facebook Messenger, WhatsApp Messenger, and YouTube, as these are commonly used platforms for communication and information exchange. We would also optimize the video formats to reduce file sizes. Finally, we would offer troubleshooting support to assist caregivers who may face technical challenges while trying to view the videos.

Enable Caregivers to Easily Share Videos With Family Members

According to caregivers, the content in the STAR-C videos had a positive impact on communication and information sharing among caregivers and other family members who assist with caregiving. For that reason, caregivers expressed a strong desire to share the video content with their extended family, including siblings and other relatives involved in caregiving responsibilities. Sharing the videos empowered caregivers to improve their caregiving practices. For example, 1 Spanish-speaking caregiver explained:

Moreover, right away, I allowed myself—I don’t know if it was allowed, but I sent that information to my sisters-in-law and they were like, “Wow, wait, so I’m going to treat my dad like this, so I can stop this. So, I got irritated by this.” My sisters said, “I got it now,” because we truly are alone, there really isn’t much of a tool at hand that leads you to something like this, and even though it might seem small, a few minutes of video, it was truly very good. So, I can’t imagine everything that’s going to come [in the future with STAR-C]. Really, I appreciate it, truly.

Another Spanish-speaking caregiver highlighted the need for education and support at the individual level. They indicated that the program’s benefits extended beyond their immediate family, potentially benefiting other caregivers in similar situations. The caregiver explained:

I believe what we were missing was educating ourselves more about this… It could be the entire family, because it was in Mexico where my mom, the rest of my siblings, and I were. We were all taking care of dad, and sometimes, even then, we couldn’t manage. We were like five adults and my dad.

In response to caregivers’ feedback, we decided that it would be appropriate to encourage Latino caregivers participating in the research study to share the STAR-C videos with other family members if they wished, regardless of whether the other family members were also enrolled in the research study. In addition, we decided to welcome other family members within the same family to enroll as study participants if they were interested and met the eligibility criteria. The program’s content could serve as a catalyst for discussion, knowledge exchange, and support among primary caregivers and their extended network of caregivers.

Theme 2: Adaptations to the “Dementia Education” Core Component

Expand Content to Improve Caregiver Understanding of Dementia

Latino family caregivers expressed the need for expanded content in the “Dementia Education” core component to enhance their understanding of dementia. They highlighted the importance of delving deeper into the nature of the disease, its progression, and the various stages it entails.
Several caregivers identified common misconceptions about Alzheimer disease and memory loss in the Latino community, expressing surprise at the late stages of dementia when physical limitations manifest. They stressed the necessity of detailed information to comprehend the evolving challenges faced by people living with dementia. Caregivers emphasized the significance of incorporating this expanded content to enhance their knowledge and foster a more holistic understanding of dementia.

But that part surprised me a lot, which is that the body itself forgets its needs. And it's the final stage. And I would like it if they did include the stages of Alzheimer's or dementia. Because for most people, if you tell them, it's like "oh, they forgot things or put something in the wrong place or they get lost." But the final stage is the one that very few people know about, it's what happens when your body, even if it gives signals, you don't recognize them. It's like a baby who can't tell you they're thirsty, they're hungry, their stomach hurts, and things like that. So, that's when it gets complicated and obviously, the end is near.

In response to caregivers’ feedback, we further expanded the content of the first video focused on teaching caregivers about ADRD. Our goal was to provide a comprehensive understanding of dementia, untangling its root causes and distinguishing it from the natural aging process. We describe the different stages of dementia to offer Latino caregivers a better understanding about types of cognitive, emotional, and behavioral changes to expect in their family member with dementia as the disease progresses.

Caregivers also expressed the need to put themselves in the shoes of the person with dementia. Further, 1 Spanish-speaking caregiver said:

The only thing is that it would also be good to mention how it's frustrating for them and for us. We should put ourselves in their shoes, how would we like to be treated if we were the ones sick, how would we want to be treated? In terms of frustration or reactions... They won't react the way we would want them to or how we would like them to react, because they have a condition, we don't.

In response to caregivers’ feedback, we integrated content into the STAR-C videos that would remind caregivers to be empathetic and to consider the perspective of people living with dementia. The added content emphasizes the importance of understanding the frustrations experienced by both caregivers and people living with dementia and the need to approach interactions with empathy and compassion.

Provide Education on how to Reduce Risk of Developing Dementia

Caregivers indicated the need for comprehensive education on reducing their own risk of developing dementia. They highlighted the importance of understanding the impact of dementia while also learning about strategies to combat it. A Spanish-speaking caregiver said:

The thing is, for example, I'm looking at my husband's case and I think about myself, and I think, "Well, what can I do to prevent what happened to my husband from happening to me?" Because what will happen if I develop dementia? What will happen to both of us? So, I would like to have more information. What can I do to avoid this? Because I am his caregiver.

Another Spanish-speaking caregiver, after learning about the prevalence of dementia among Hispanics in the STAR-C video, expressed the need for more information on prevention and early signs. They said:

That's really good [the information in the module]. I was genuinely surprised when it said that one in every eight Hispanics has or will develop, right? So, it's a bit alarming, and I think, wow, I don't know, I would like to know more about whether there would be any way to prevent it. It would be great to have more information about prevention or the signs—as it says there, some forgetfulness is normal, certain forgetfulness, right? Like now, being busy with a thousand things, I forgot, and believe me, it happens to me, but I know it's because I have a lot on my plate and I try to do them all. But I would really like if there was information about whether there're any way to prevent this condition.

Caregivers’ feedback confirmed our previous findings about the need to modify the STAR-C content to include...
comprehensive education on dementia. In response to caregivers’ feedback, we will also be providing caregivers with information on reducing their own risk of developing the condition. By providing caregivers with guidance and knowledge on prevention strategies, STAR-C can empower them to take proactive steps in safeguarding their cognitive health and that of their loved ones.

**Theme 3: Adaptations to the “ABC Problem-Solving” Core Component**

**Simplify the “ABC” Problem-Solving Acronym in Spanish**

The “ABC” problem-solving approach in the STAR-C program was regarded as helpful by some Latino caregivers, who found the provided examples to be effective and relatable. However, it was acknowledged that understanding the acronym (“Activators, Behaviors, Consequences”) could be a bit challenging for others. Further, 1 Spanish-speaking caregiver explained why she liked the ABC problem-solving approach:

> So, I found the video to be very original, very realistic. It was done very well because I felt identified. The three ways they presented it, in A, B, and C, personally, it felt very real to me. I loved it, almost perfect, because these are situations that do happen and changes that we do need to make. From the beginning, we don’t know how to do it, but with this video or the app that they’re going to develop, it seems very practical to me because it will provide a lot of tools and strategies to people who have no idea how to go about it. Like us in the beginning, we were learning as we went through each day.

Another Spanish-speaking caregiver initially faced challenges in understanding the ABC acronym but gained clarity once the video explained it further.

> Yes, some of those words were a bit difficult for me to understand, but later on, it was explained what each of them meant.

Further, 1 Spanish-speaking caregiver reported that while they could understand the content, including the ABC acronym, well due to their extensive Spanish language skills, they acknowledged that a person with limited education might struggle to understand it clearly.

> I can understand it perfectly, but I think I can understand because I have a very good Spanish. I was a Spanish teacher for many years, so my language and vocabulary are quite extensive. However, I believe that if the same video had to be heard by someone with limited education, they probably wouldn’t understand it, at least not clearly.

These quotes highlight the caregivers’ perspective on the need to simplify the “ABC” problem-solving acronym in Spanish. While some caregivers found the explanations of each letter in the acronym helpful, others reported that individuals with limited education may potentially face challenges in understanding the acronym. In response to caregivers’ feedback, we simplified the ABC acronym to enhance comprehension and accessibility for a wider range of caregivers. Initially, the ABC acronym was translated as A for “activadores” (activators), B for “comportamiento” (behavior), and C for “consecuencias” (consequences). However, based on the feedback received, we took an additional step to simplify it entirely in Spanish, resulting in the revised form as “las 3 Cs” (the 3 Cs) representing C for “causas” (causes), C for “comportamiento” (behavior), and C for “consecuencias” (consequences).

**Add More Problem-Solving Examples With Different BPSDs**

Caregivers conveyed a strong desire for the STAR-C program to incorporate a greater variety of problem-solving examples that cover different BPSDs. While some caregivers appreciated the existing examples, they emphasized the importance of including a more extensive range of stories and behaviors to address the diverse challenges encountered in dementia care. For example, 1 Spanish-speaking caregiver stated:

> I liked everything, the only thing is that I would like them to add a bit more different stories, with different behaviors.

Another Spanish-speaking caregiver shared their personal experience with their mother’s behavior and the importance of addressing such situations.

> Yes, my mom experiences a lot of panic episodes, and I didn’t see that in the video, so in my own way, I handled panic situations in her illness, like waking up at night in a panic: “Where am I? Who am I? Where are we?” So, I didn’t see in the video strong things like that, like screaming, situations where you don’t know what to do as a family member, so you just hug them. I hugged my mom, I hugged her, I hugged her, and I said, “Calm down, we’re okay. I’m your daughter, we’re here.” I mentioned the house, the surroundings, everything. But she has those night panics out of fear very often.

These quotes underscore the caregivers’ interest in having a more extensive selection of problem-solving examples that address various BPSDs, such as aggression and panic attacks. In response to caregiver’s feedback, we will include a broader range of scenarios and behaviors, so that caregivers can gain invaluable insights and strategies to effectively manage the diverse challenges associated with BPSDs.

**Demonstrate Problem-Solving With Real People**

Latino caregivers reported a desire for more realistic and relatable problem-solving examples in the STAR-C program. They suggested incorporating videos featuring real interactions between caregivers and persons living with dementia to enhance the learning experience. Further, 1 Spanish-speaking caregiver explained:

> Perhaps, in the example—I mean, I don't know how much of the video or program is left [to be developed], but maybe, I don't know, perhaps the depiction of two people acting out the situation, maybe it would look much more professional or
Another Spanish-speaking caregiver reported interest in observing caregiving interactions to better understand effective communication strategies.

Perhaps some of this information would be clearer in a video format or as an interaction between two people. Yes, specifically with an Alzheimer’s patient. Because if they're [the program] going to be incorporating strategies on how to talk, how to interact, it would be very useful to see the behavior live; how it's happening, how one can communicate with that person.

Incorporating videos of real people and interactive caregiving scenarios in the STAR-C program can provide caregivers with tangible and relatable examples of problem-solving techniques. Based on caregivers’ feedback, future adaptations of STAR-C should demonstrate the app of the problem-solving techniques with real people. These demonstrations may enhance caregivers’ experience with the program by presenting solutions in a dynamic and engaging manner.

**Theme 4: Adaptations to the “Caregiver Support Strategies” Core Component**

Latino caregivers emphasized the importance of seeking assistance and support to alleviate the burdens of caregiving. Many highlighted the value of reaching out to family members or friends when feeling fatigued or overwhelmed, recognizing the need for rest and self-care. A Spanish-speaking caregiver described how the STAR-C program needed to encourage caregivers to take care of their own health and well-being. The caregiver said:

>[For caregivers] to be well-rested. That's why I mention seeking help in some way, even from family or friends, you can say “I need—” when you feel tired or stressed, very stressed, you need to be able to call someone and say, “Can you take care of my mom for a while? Can you take care of my dad or my wife? I need a break.” Because if you're not well-rested or already feeling overwhelmed—I say this because, for example, my sister, when she was taking care of my mom, my sister wasn't emotionally well. She had her own problems, so when it was her turn to care for my mom, there were almost always issues because she didn't have enough patience.

Participants collectively acknowledged the importance of supporting caregivers themselves, not just the care recipients. They emphasized the need for caregivers to take regular breaks and engage in activities that promote mental well-being. For example, a Spanish-speaking caregiver stated:

>**I just wanted to mention that also the people—sometimes the people who care for others also need to be taken care of, to take a break or do something different in order to continue, to be mentally well enough to keep taking care of our loved ones.**

This comment highlights the importance of recognizing the caregiver’s mental and emotional health as it directly impacts their ability to provide effective care. In response to caregivers’ feedback, we integrated content throughout the various modules (not just the last module) about the importance of caregiver health and well-being. The content emphasizes the need for caregiver rest, support, and self-care, so that caregivers can be better equipped to provide optimal care to their loved ones.

**Discussion**

**Principal Results**

This study’s objective was to test and refine the preliminary cultural adaptations to the STAR-C web-based training modules for Latino caregivers. Our qualitative analysis identified key adaptations required in (1) the delivery of STAR-C, (2) the “dementia education” core component, (3) the “ABC problem-solving” core component, and (4) the “caregiver support strategies” core component. Caregivers expressed a desire for STAR-C videos to be accessible through audio-only formats, and they highlighted the importance of making the videos available on various platforms, including those commonly used in their community. They also emphasized the need for the videos to be easily shareable with family members to enhance communication and caregiving practices. In response, we tailored the narration script to facilitate audio-only access and optimized video formats for widespread accessibility. Moreover, caregivers requested comprehensive education on dementia prevention and understanding the disease’s stages, advocating for expanded content within the “Dementia Education” core component. This resulted in additional content aimed at enhancing caregivers’ understanding of dementia progression and challenges. To further foster empathy, caregivers suggested highlighting the importance of viewing the world from the perspective of individuals living with dementia. Consequently, we incorporated elements emphasizing empathy into the program. Caregivers also sought to incorporate strategies for reducing the risk of developing dementia, leading to the inclusion of content addressing preventive measures. Within the “ABC Problem-Solving” core component, caregivers expressed a need to simplify the “ABC” acronym in Spanish and expand problem-solving examples. We responded by simplifying the acronym and incorporating more problem-solving scenarios to cater to diverse challenges. Additionally, caregivers desired realistic problem-solving examples featuring real interactions between caregivers and persons with dementia. Lastly, caregivers emphasized the importance of caregiver well-being, prompting us to underscore self-care, rest, and seeking support to ensure caregivers’ mental and emotional health is prioritized. These adaptations collectively refine the STAR-C web-based training modules for Latino caregivers and provide a culturally tailored, evidence-based intervention to support dementia caregiving in this community.

**Comparison With Prior Work**

Our study identified several key adaptations needed to enhance the cultural relevance and effectiveness of the STAR-C web-based training modules for Latino caregivers. Notably, caregivers expressed a strong desire for videos to be accessible through audio-only formats, a finding that aligns with a study
conducted among Hispanic participants from Spain [19], where a similar preference for audio accessibility was observed. The importance of making videos available on various platforms, as highlighted in our study, is also consistent with research where caregivers expressed openness to learning caregiving information from diverse sources and settings [20].

The cultural adaptations identified in this study to enhance the relevance and effectiveness of the STAR-C web-based training modules for Latino caregivers may have applicability to caregivers from other racial and ethnic minority groups as well. For example, collectivist and familialism values tend to be strong in Asian cultures that emphasize family and community interdependence [21,22]. As such, making the STAR-C videos easily shareable with family members to facilitate communication and coordinated caregiving practices may also resonate with Asian caregivers.

Furthermore, our study aligns with prior research by emphasizing the need for comprehensive education on dementia prevention and understanding the disease’s stages [20,23]. In a study conducted among various ethnic groups, including Hispanic or Latinos, African Americans, and Asian Americans, it was found that a common lack of knowledge about the early signs of Alzheimer disease existed [23]. This knowledge gap underscores the importance for comprehensive dementia education, a point that resonates with our study’s findings. Additionally, another study similarly stressed the importance of greater education regarding the diversity and spectrum of dementia-related symptoms [20].

Caregivers also strongly advocated for simplifying the ABC acronym in Spanish, adding more problem-solving examples, and demonstrating problem-solving with real individuals. These adaptations align with a study assessing Spanish language health information via videos [24], which found that participants preferred videos featuring increased actor participation, as it enhanced their ability to relate the content to real-life situations.

Limitations

Our study has limitations worth noting. First, the low-fidelity prototypes of the STAR-C web-based training modules (i.e., videos) we used in this study lacked the realistic appearance and comprehensiveness of the full virtual STAR-C program. This may have limited our ability to obtain accurate perceptions and feedback, as caregivers might not have fully understood or engaged with the low-fidelity prototypes as they would have with a higher-fidelity version of the full STAR-C program. As a next step in this research, we plan to pilot test high-fidelity prototypes of the STAR-C web-based training modules. Second, caregivers’ feedback about STAR-C is influenced by their personal preferences and experiences, which may result in us inadvertently prioritizing features or solutions that appeal to the specific subgroup of Latino caregivers participating in this study rather than broader populations of Latino caregivers. We tried to address this limitation by interviewing a diverse group of caregivers including both male and female caregivers (most Latino caregivers tend to be women), caregivers of different ages and family roles (i.e., spouse vs child caregiver), and caregivers with different educational backgrounds.

Conclusions

This study addresses a critical gap in the field of dementia caregiving, particularly for Latino caregivers who face unique challenges and disparities in health outcomes. The culturally adapted version of the virtual STAR-C program represents a significant step forward in bridging this gap. Our findings highlight the importance of tailoring interventions to meet the specific needs of Latino caregivers, considering cultural values and beliefs that shape caregiving interactions and experiences. The key adaptations we needed to make to the STAR-C web-based training modules, including accessibility improvements, expanded dementia education, enhanced problem-solving strategies, and a focus on caregiver well-being, underscore the importance of cultural adaptation of evidence-based caregiver interventions for Latino families. While further research is needed to assess the efficacy of these adaptations, we believe that the virtual STAR-C intervention has the potential to improve health care services and health outcomes for Latino people living with dementia and their family caregivers. By addressing the pressing need for culturally adapted evidence-based interventions, we aim to promote more favorable experiences with the intervention and ultimately reduce the health disparities associated with dementia caregiving in this community. This work contributes to the broader mission of advancing health care services for older adults by using technological innovations, serving the interest of health professionals and family caregivers of older adults.

Acknowledgments

This study was funded by the Alzheimer’s Association (AARG-20-685260, principal investigator: MR).

Conflicts of Interest

None declared.

References


Abbreviations

- **ABC**: Activators-Behaviors-Consequences
- **ADRD**: Alzheimer disease and related dementias
- **BPSD**: behavioral and psychological symptoms of dementia
- **REDCap**: Research Electronic Data Capture
- **STAR-C**: STAR caregiver

Edited by Y Jiang; submitted 16.10.23; peer-reviewed by F Medina, I Wilson, SQ Yoong; comments to author 15.11.23; revised version received 06.12.23; accepted 06.12.23; published 11.01.24.

Please cite as:

Garcia CN, Duran MC, Ramirez M
Refining Cultural Adaptations of a Behavioral Intervention for Latino Caregivers of People Living With Dementia: Qualitative Interview Study in Washington State
JMIR Aging 2024;7:e53671
URL: https://aging.jmir.org/2024/1/e53671
doi:10.2196/53671
PMID:38206663

©Celeste N Garcia, Miriana C Duran, Magaly Ramirez. Originally published in JMIR Aging (https://aging.jmir.org), 11.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Expectation, Attitude, and Barriers to Receiving Telehomecare Among Caregivers of Homebound or Bedridden Older Adults: Qualitative Study

Pansiree Onseng¹, MD; Wichuda Jiraporncharoen¹,², MSc, MD; Sasiwimon Moonkayaow¹, MD; Pimchai Veerasirikul¹; Nutchar Wiwat kunupakarn¹,², MD; Chaisiri Angkurawaranon¹,², MD, PhD; Kanokporn Pinyopornpanish¹,², MD

¹Department of Family Medicine, Chiang Mai University, Chiang Mai, Thailand
²Global Health and Chronic Conditions Research Group, Chiang Mai University, Chiang Mai, Thailand

Abstract

Background: In recent years, telehomecare has become an increasingly important option for health care providers to deliver continuous care to their patients.

Objective: This study aims to explore the expectations, attitudes, and barriers to telehomecare among caregivers of homebound or bedridden older adults.

Methods: This qualitative study used semistructured interviews to explore caregivers’ perspectives on telehomecare for homebound or bedridden older adults. The study adhered to the SRQR (Standards for Reporting Qualitative Research) guidelines. Participants were selected using convenience sampling from caregivers of homebound or bedridden older adults with experience in both in-person home visits and telehomecare services provided by the Department of Family Medicine at Chiang Mai University, in an urban area of Chiang Mai Province in Northern Thailand. Semistructured interviews were conducted. The interviews were audio recorded with participant consent and transcribed verbatim. The framework method was used, involving multiple readings of transcripts to facilitate familiarization and accuracy checking. The study used the technology acceptance model and comprehensive geriatric assessment as the analytical framework.

Results: The study included 20 caregivers of older adult patients. The patients were predominantly female (15/20, 75%), with an average age of 86.2 years. Of these patients, 40% (n=8) of patients were bedridden, and 60% (n=12) of patients were homebound. Caregivers expressed generally positive attitudes toward telehomecare. They considered it valuable for overall health assessment, despite recognizing certain limitations, particularly in physical assessments. Psychological assessments were perceived as equally effective. While in-person visits offered more extensive environmental assessments, caregivers found ways to make telehomecare effective. Telehomecare facilitated multidisciplinary care, enabling communication with specialists. Caregivers play a key role in care planning and adherence. Challenges included communication issues due to low volume, patient inattention, and faulty devices and internet signals. Some caregivers helped overcome these barriers. The loss of information was mitigated by modifying signaling equipment. Technology use was a challenge for some older adult caregivers. Despite these challenges, telehomecare offered advantages in remote communication and resolving scheduling conflicts. Caregivers varied in their preferences. Some preferred in-person visits for a broader view, while others favored telehomecare for its convenience. Some had no strong preference, appreciating both methods, while others considered the situation and patient conditions when choosing between them. Increased experience with telehomecare led to more confidence in its use.

Conclusions: Caregivers have positive attitudes and high expectations for telehomecare services. Although there may be barriers to receiving care through this mode, caregivers have demonstrated the ability to overcome these challenges, which has strengthened
their confidence in telehomecare. However, it is important to enhance the skills of caregivers and health care teams to overcome barriers and optimize the use of telehomecare.

(JMIR Aging 2024;7:e48132) doi:10.2196/48132

KEYWORDS

telehomecare; telemedicine; telehealth; caregivers; older adults; attitudes

Introduction

Thai society has come to recognize the challenge of the aging society and that there is a rapid growth in the number of homebound or bedridden older adult patients with chronic diseases [1]. Homebound refers to a state in which one’s living space is limited to their residence, yet they retain the ability to move within their home. On the other hand, being bedridden is a condition in which a person has experienced significant physical weakness and can no longer move about freely, thus being confined to their bed. These two statuses are typically caused by the patient’s comorbidities and often result in the need for comprehensive care and caregiver [2-4]. As a result, coordinated health care services are essential for these individuals, and home health care services have been found to be effective in improving patient outcomes. A long-term care policy has been implemented to provide home care and social support for this population [1,5]. These services are particularly useful for doctors to assess a patient’s living conditions and have been shown to decrease hospitalization, improve physical and psychosocial health, and enhance the overall quality of life for older adults [6,7].

Telemedicine has emerged as a potential solution to bridge the gap between homebound or bedridden patients and health care services [8]. By allowing health care providers to connect with patients remotely, telemedicine can reduce travel costs, provide convenience, and help control the spread of the pandemic [9]. However, there are also disadvantages and challenges associated with this technology, such as the lack of face-to-face communication between doctors and patients and the need for patients to have the technical skills to use these devices [10]. These factors may make the home visit system unsustainable [10].

Over the past 3 years, the COVID-19 pandemic has further highlighted the need for telemedicine, as in-person home visits have been interrupted and transformed into telehomecare for patients requiring continuous health care at home [11-13]. Telehomecare is a form of telemedicine that combines videoconferencing and health monitoring between homebound or bedridden patients and home health care providers [14]. In order to conduct telehomecare for homebound or bedridden older adult care, the caregiver has the main role of helping monitor the patient and to communicate all information to health care team. Therefore, if the caregivers are also older adults and are not proficient with technology, this can pose a barrier, especially for patients living in remote areas [15].

Prior studies on the perception or perspective of the use of telemedicine in caring for older adults have primarily focused on the viewpoints of physicians [16,17] and older adult patients [18,19], rather than caregivers. Physicians have also reported perceiving advantages in the use of telemedicine for older adults, including the reduction of postponed care, the promotion of timely care, enhanced physician efficiency, improved communication with both patients and caregivers, the alleviation of patient travel burdens, and the facilitation of health outreach and educational efforts [17]. The use of technology for health care in older adults could be influenced by their age and the complexity of diseases they face [19], impacting their acceptance of technology [20] and use behavior [21]. It has been noted that the technological acceptance of patients and providers appears to have positive effects on patient outcomes, including self-management and readmissions. The literature also indicates that family caregivers play a vital role in assisting the patient’s decision to adopt and engage with technology [22]. Physicians are aware that caregivers have a role in deciding whether to use telehealth services [16]. Overall, it is important to assess caregivers’ genuine perspectives and suggestions.

A recent study from the United States has explored family caregivers’ perceived benefits and barriers to telemedicine visits for community-dwelling older adults with chronic diseases [23]. Caregivers reported effectiveness in increasing access and continuity of care, but they also expressed concerns about their older relatives’ ability to access and use the technology independently and difficulty in establishing patient-provider rapport. However, these results could be different in Thailand and among families of older adults with more limited capacity, such as the homebound or bedridden population. Thus, this study aims to investigate the expectations, attitudes, and barriers concerning telehomecare among caregivers of homebound or bedridden older adults. Specifically, it will explore caregivers’ expectations regarding the perceived usefulness of the service in delivering comprehensive care for older adults with limited mobility; the barriers that affect the perceived ease of using the service; and their overall attitudes toward the care provided, including their preferences. The results of this study will provide useful insights for the creation of effective care plans for homebound or bedridden older adult patients and their caregivers.

Methods

Study Design

This study was a qualitative study, using semistructured interviews to investigate the perspectives of using telehomecare for homebound or bedridden older adults, as perceived by their caregivers. The study was reported according to the SRQR (Standards for Reporting Qualitative Research) guidelines [24].

Ethical Considerations

Ethical approval was obtained through the Medical Ethical Committee of Chiang Mai University, Chiang Mai, Thailand.
whether telehomecare is suitable for providing comprehensive

Data Collection

Semistructured in-depth interviews were conducted between November 2021 and March 2022. There were two parts to data collection as follows: (1) participant characteristics and (2) in-depth interviews about the perspectives of using telehomecare for homebound or bedridden older adults. The patient interviews were conducted in Thai by a single researcher (PO) who was trained in the interview method and interview questions by the qualitative researchers (KP and WJ). Each interview was conducted in a private counseling room to help ensure comfort and confidentiality. Each interview was audio recorded with the participant’s consent, followed by verbatim transcription. PO and KP subsequently reviewed the audio record, and any issues that needed further exploration were discussed. Participant characteristics including sex, age, occupation, relationship with the patient, and the duration of care as the caregiver were recorded. Patient diagnoses and homebound or bedridden status were also gathered. Open-ended interview guides were designed and then piloted to make sure that participants were allowed to share their thoughts freely. The interview guide was described in Textbox 1.

Textbox 1. The interview guide.

- How long has your family been receiving home care and telehomecare?
- What was the main reason for switching from in-person care to telehomecare?
- Do you perceive telehomecare as useful for your family?
- As a caregiver for an older adult patient, do you feel confident about receiving care for the patient through telehomecare?
- Can you compare the effectiveness of care between in-person home visits and telehomecare for older adult patients?
- Are you satisfied with these services? Which one do you prefer and why?
- What do you see as potential barriers or concerns regarding the use of telehomecare for caring for older adult patients?

Data Analysis

We followed the framework method for the analysis of qualitative data [26]. Each transcript was read multiple times by the 2 researchers (PO and KP) to aid familiarization and to check the accuracy of each transcript. The key elements of the technology acceptance model and comprehensive geriatric assessment for older adult patient care were used as the analytical framework to help identify key issues, concepts, and themes. The study aimed to address the 2 main research questions, which align with the technology acceptance model in the context of providing care for older adults from the perspective of caregivers. The technology acceptance model is the theory that suggests the predictors toward the acceptance and rejection to use technology. These predictors include (1) perceived usefulness and (2) perceived ease of use which are influenced by, for example, technology anxiety or experience, subjective norms, expectations, trust, cultural diversities, and technology characteristics [27]. These factors, in turn, shape attitudes toward using telehomecare services.

The first question focused on perceived usefulness (expectation), incorporating comprehensive geriatric assessment to determine whether telehomecare is suitable for providing comprehensive home-based care for older adults. Comprehensive geriatric assessment is usually carried out while providing care for older adults at home to increase the understanding of an older person’s care needs and preferences and to help in finding common goals of care [28,29]. Its three key elements include (1) comprehensiveness (physical, psychological, and environmental issues); (2) multidisciplinarity; and (3) person-centered goal setting. The second question concentrated on the perceived ease of use of telehomecare for caregivers caring for older adults. This encompassed concerns and perceived barriers.

To explore our research questions, we adopted the framework method, which used a combined approach, incorporating both deductive and inductive elements for the analysis [26]. The deductive approach was applied to the preexisting theories, specifically the technology acceptance model and comprehensive geriatric assessment. Meanwhile, the inductive approach was used to analyze the expectations, attitudes, and barriers associated with telehomecare among caregivers of homebound or bedridden older adults.

The researchers compared the identified codes and discussed the similarities and differences until they reached a consensus on the emergent themes and subthemes. Then, the preliminary

Recruitment

Participants were selected using convenience sampling from caregivers of homebound or bedridden older adult patients who had experience with both in-person home visits and telehomecare services provided by the Department of Family Medicine at Chiang Mai University. The samples were selected from a list of families of older adults who had received care within the past 2 months. Research assistants reached out to these families via phone calls until the total expected sample size was achieved. The study was conducted in an urban area of Chiang Mai Province in Northern Thailand. Inclusion criteria were that participants must be the primary caregiver and speak Thai, while caregivers with communication difficulties were excluded. The researchers expected to recruit at least 20 participants to achieve data saturation, based on similar qualitative studies [25].

Informed consent was obtained from all participants. The privacy and confidentiality of human participants were rigorously protected throughout this study. Data were either anonymized or deidentified, as applicable, to safeguard the identity of participants. Participants in the study received a monetary compensation of 100 (US $2.7) for their participation.

Data Collection

Semistructured in-depth interviews were conducted between November 2021 and March 2022. There were two parts to data collection as follows: (1) participant characteristics and (2) in-depth interviews about the perspectives of using telehomecare for homebound or bedridden older adults. The patient interviews were conducted in Thai by a single researcher (PO) who was trained in the interview method and interview questions by the qualitative researchers (KP and WJ). Each interview was conducted in a private counseling room to help ensure comfort and confidentiality. Each interview was audio recorded with the participant’s consent, followed by verbatim transcription. PO and KP subsequently reviewed the audio record, and any issues that needed further exploration were discussed. Participant characteristics including sex, age, occupation, relationship with the patient, and the duration of care as the caregiver were recorded. Patient diagnoses and homebound or bedridden status were also gathered. Open-ended interview guides were designed and then piloted to make sure that participants were allowed to share their thoughts freely. The interview guide was described in Textbox 1.

Textbox 1. The interview guide.

- How long has your family been receiving home care and telehomecare?
- What was the main reason for switching from in-person care to telehomecare?
- Do you perceive telehomecare as useful for your family?
- As a caregiver for an older adult patient, do you feel confident about receiving care for the patient through telehomecare?
- Can you compare the effectiveness of care between in-person home visits and telehomecare for older adult patients?
- Are you satisfied with these services? Which one do you prefer and why?
- What do you see as potential barriers or concerns regarding the use of telehomecare for caring for older adult patients?

Data Analysis

We followed the framework method for the analysis of qualitative data [26]. Each transcript was read multiple times by the 2 researchers (PO and KP) to aid familiarization and to check the accuracy of each transcript. The key elements of the technology acceptance model and comprehensive geriatric assessment for older adult patient care were used as the analytical framework to help identify key issues, concepts, and themes. The study aimed to address the 2 main research questions, which align with the technology acceptance model in the context of providing care for older adults from the perspective of caregivers. The technology acceptance model is the theory that suggests the predictors toward the acceptance and rejection to use technology. These predictors include (1) perceived usefulness and (2) perceived ease of use which are influenced by, for example, technology anxiety or experience, subjective norms, expectations, trust, cultural diversities, and technology characteristics [27]. These factors, in turn, shape attitudes toward using telehomecare services.

The first question focused on perceived usefulness (expectation), incorporating comprehensive geriatric assessment to determine whether telehomecare is suitable for providing comprehensive home-based care for older adults. Comprehensive geriatric assessment is usually carried out while providing care for older adults at home to increase the understanding of an older person’s care needs and preferences and to help in finding common goals of care [28,29]. Its three key elements include (1) comprehensiveness (physical, psychological, and environmental issues); (2) multidisciplinarity; and (3) person-centered goal setting. The second question concentrated on the perceived ease of use of telehomecare for caregivers caring for older adults. This encompassed concerns and perceived barriers.

To explore our research questions, we adopted the framework method, which used a combined approach, incorporating both deductive and inductive elements for the analysis [26]. The deductive approach was applied to the preexisting theories, specifically the technology acceptance model and comprehensive geriatric assessment. Meanwhile, the inductive approach was used to analyze the expectations, attitudes, and barriers associated with telehomecare among caregivers of homebound or bedridden older adults.

The researchers compared the identified codes and discussed the similarities and differences until they reached a consensus on the emergent themes and subthemes. Then, the preliminary
Sociodemographic Data of Caregivers

A total of 20 caregivers were recruited, of which 25% (n=5) were male and 75% (n=15) were female, with an average age of 59.5 (SD 14.27) years. Fourteen (70%) of the caregivers were identified as the children of the patients, 15 (75%) caregivers had at least a bachelor’s degree or equivalent level of education, and the average duration of patient care was 3.3 (SD 2.9) years. Table 1 provides a summary of the general information about the caregivers. The average interview duration was 28.45 minutes, ranging from 20 to 47 minutes.

### Table 1. Sociodemographic data of caregivers (N=20).

<table>
<thead>
<tr>
<th>Item</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5 (25)</td>
</tr>
<tr>
<td>Female</td>
<td>15 (75)</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>59.5 (14.27)</td>
</tr>
<tr>
<td>Range</td>
<td>26-87</td>
</tr>
<tr>
<td><strong>Relation to patient, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Spouse</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Siblings</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Child</td>
<td>14 (70)</td>
</tr>
<tr>
<td>Grandchild</td>
<td>3 (15)</td>
</tr>
<tr>
<td><strong>Highest education, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Primary school</td>
<td>2 (5)</td>
</tr>
<tr>
<td>Secondary school</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Bachelor’s degree and higher</td>
<td>15 (75)</td>
</tr>
<tr>
<td><strong>Duration of receiving home health care (years)</strong></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>3.3 (2.9)</td>
</tr>
<tr>
<td>Range</td>
<td>1.5-15</td>
</tr>
</tbody>
</table>

Health Status of Older Adult Patients

The majority (15/20, 75%) of patients were female, with an average age of 86.2 (range 66-95) years. Among these patients, 8 (40%) patients were bedridden and 12 (60%) patients were homebound. The most common comorbidities included hypertension (15/20, 75%), dyslipidemia (7/20, 35%), dementia (7/20, 35%), and stroke (6/20, 30%). Additionally, 45% (n=9) of patients reported needing assistance for using technology, while 40% (n=8) of patients indicated that they were unable to use technology.

Caregivers’ Expectations, Attitudes, and Barriers Toward Telehomecare

Overview

Based on the interviews, it was found that caregivers of older adult patients have specific attitudes and expectations toward taking care of the older adult by using telehomecare. They see it as a valuable service for overall health assessment, although they acknowledge limitations in certain aspects. Communication and device usability challenges were noted, but there is a clear advantage to telehomecare in specific cases compared with in-person visits. Confidence in this service is widespread, yet individual preferences are influenced by family circumstances and situations. Multimedia Appendix 1 shows the framework matrix. The details are provided in the following sections.

It is a Service That Provides an Overall Assessment of a Patient's Health, but Some Aspects May Be Limited

When surveying caregivers’ perceived usefulness toward telehomecare for older adult patients within the comprehensive geriatric assessment framework, differing attitudes emerged. Overall, caregivers displayed a favorable attitude toward telehomecare, despite some limitations in specific areas.
Physical Assessment

With regard to the physical assessment of the patients, telehomecare can be more limited than in-person home visits. However, caregivers can mitigate this by providing clearer information through photos and videos.

- Similar to the case of my mother’s mole which turned into a malignant tumor, it was not noticed clearly through VDO call. [Participant 019]
- It doesn’t affect me. If grandma has a pain or where there is any wound, the doctor will ask to take pictures or ask her to walk around to observe the symptom through video call. [Participant 014]

Psychological Assessment

Caregivers believe that psychological assessments yield similar outcomes through both telehomecare and in-person visits, as doctors can ask questions and provide privacy for patients during an assessment session.

- It’s not different. The doctor asks the same questions, such as “Are you worried about anything?” or “Would you like me to help you with anything?” [Participant 017]

Environmental Assessment

In-person visits offer a broader environmental assessment experience compared with telehomecare, leading to more precise advice on home environment modification. However, caregivers have found ways to use telehomecare effectively, such as walking around with the device.

- If the doctor was to visit in person, they would be able to see the home environment and provide advice on how to improve it. If it is stuffy, smelly, unsightly, not suitable for the patient, then the doctor can suggest more than a VDO call. This is because sometimes we don’t show the home environment as much. [Participant 013]

Multidisciplinary Care

Telehomecare facilitates engagement with specialists such as dieticians and physiotherapists, thus enhancing patient care.

- The telehomecare allows us to have more knowledge from other members of the health care team, especially about how to do physical therapy and how to manage the diet for the patient. [Participant 015]

Goal of Care and Care Planning

Caregivers play a key role in facilitating communication and treatment plan adherence during telehomecare, ensuring patients follow prescribed plans effectively.

- We’re able to keep up with the treatment plan even during telehomecare because when the doctor has a video call with my grandma, I always listen and take notes, and then I go over what the doctor has told her. I take notes of it all. [Participant 002]

Communication and Device and Usability Limitations Are Also Challenges

In telehomecare, despite the content, care processes, and follow-ups remaining the same, challenges still arise. The main barriers are related to communication and issues with the equipment, but most caregivers are capable of handling these challenges.

Communication

Challenges include low volume from either the patient or health care worker, low patient inattention, and faulty devices, leading to unclear communication. Caregiver presence during sessions can help overcome these barriers.

- Grandma’s voice is quite soft. She can’t speak loud making it difficult for the doctor to hear [Participant 002]
- During telehomecare, if mother is sitting and trying to listen, it may be necessary for a caregiver to be present to help explain things, as she may have a reduced ability to listen and understand due to her age and weakening internal systems. [Participant 015]

Technologies and Signaling Systems

Loss of information may occur in telehomecare, but some caregivers have modified or changed the signaling equipment and environment to enhance it, resolving this issue.

- The sound is a bit lacking, but it's probably my Internet. After changing the Wifi, it’s better. My mother’s room lacks internet signal. [Participant 007]
- Barriers? There are some when we use VDO call. Because the phone is old, and we changed the area where we put the phone, so we received a call late. We don’t use Wifi. [Participant 018]

Ability to Use Technology

Older caregivers may encounter difficulties due to their unfamiliarity with technology, affecting their ability to use telehomecare equipment.

- I am not good at using the smartphone. Sometimes they ask me to take pictures of this and that area and I don’t really know how to shoot and am not very proficient in using the device. [Participant 013]

There is an Obvious Advantage of Telehomecare in Limited Cases Rather Than In-Person Home Visits

Despite some challenges and difficulties, telehomecare offers a clear advantage by providing a means of communication in remote areas.

Scheduling Conflicts

Telehomecare can resolve scheduling conflicts caused by caregiver appointments, enabling remote participation when in-person visits are not possible.

- There are times when the older adult patient is not at home during the appointment time. Recently, Grandma was at the center. But we communicated through Line in this group. And then we turn on the
Because the last time the doctor and the nurse came to visit the house, they gave me notice and later, mother had discovered a mole that protruded into a malignant tumor; we didn’t know it. We thought it was a normal wart. Here, in-person home visits are very helpful because sometimes caregiver doesn’t know what it is. [Participant 019]

Preference of Telehomecare

Two caregivers prefer telehomecare over in-person visits because it is more convenient as they require less preparation. Often, during home visits, the patient may be sleeping and most of the communication is with the caregiver.

She prefers a VDO call because it is convenient for both doctor and patient. Sometimes the patient will be taking a nap. Most of the patients are old as grandma start napping a lot. If the doctor comes and grandma is not talking to him anymore, she’ll take a nap. This is a waste of the doctor’s time. [Participant 016]

Preference for Both Methods

Five caregivers had similar preferences for both forms of services. They appreciated that both approaches provided an equivalent level of service and treatment, meeting their expectation and goals of health care services during examinations and treatment processes.

I like both. I had no problems with both of them, Through VDO calls, we also get what we need for patient care - medicines and lab results. For me, I don’t have a preference for either method because the specific advantages are different. [Participant 003]

Additionally, some of the caregivers had no preference between in-person home visits and telehomecare as they found advantages in both methods depending on the situation, such as the current pandemic situation and the patient’s condition.

It can be both. But now the doctor says that during the COVID pandemic, he still needs to communicate through VDO call as it is good for both sides. As for the in-person home visit, it provides a visual interaction, but during A VDO call, we still have talking, interviewing, asking questions, providing the health care process which is similar kind of work as well. [Participant 014]

Discussion

Principal Findings

The study discovered that telehomecare is not the same as in-person visits. It may present certain difficulties and barriers, particularly in the aspects of assessment, interaction, and technology use. Despite these barriers, it has the potential to fulfill requirements and bolster trust among those caregivers who provide care to older adults at home to a similar extent as a face-to-face visit. The majority of patients prefer both forms
of care, but it is important to adapt use to fit specific circumstances and conditions.

**Comparison With Prior Work**

The caregivers of homebound or bedridden older adults expressed positive attitudes toward telehomecare and are prepared to familiarize themselves with this service, comprehending its limitations and the need for its use instead of in-person visits. The previous scope review for patients receiving the telehomecare service at home found that patients desire telehomecare as a supplementary channel, offering easier access to health services from the health care team. This can enhance symptom management and encourage greater self-care [30]. Additionally, 1 study shows that telehomecare patients have greater expectations of telehomecare for the quality and efficiency of patient care, a positive relationship with the team, reduced travel costs [31], and a sense of continuous health monitoring and reassurance from their health care team [32]. If these concerns are addressed in patient care, it will heighten the confidence and satisfaction of both patients and their families.

Moreover, previous studies have found that the majority of participants with a negative attitude toward telehomecare services need more face-to-face health checks than those monitored by telehomecare. This is because they did not have a sense of touch as in in-person visits and were not familiar with using devices. This caused barriers. Conversely, those who are already comfortable with the technology tend to have a more favorable view and see telehomecare as a convenient option, eliminating the need for physically transporting patients to a hospital [8]. These findings align with the results of this study.

**Recommendations for Enhancing Telehomecare Services**

Based on these results, there are 4 recommendations for enhancing telehomecare services as provided in the following sections.

**The Provider Should Enhance the Patients’ or Caregivers’ Skills to Do Self-Health Assessment**

Our study results revealed that it is possible to conduct an overall assessment of a patient’s health through telehomecare, but there may be limitations in assessing certain aspects, such as physical assessment. Telehomecare’s limitations make it hard to conduct comprehensive videoconferencing or telephone-based health checks, leading to potential medical errors. To ensure effective patient care, family members and caregivers must be involved [33]. Educating them on symptom recognition and initial assessments can help detect abnormalities early, leading to prompt notification for further assessment by the health care team [23]. It is important to provide training and education to caregivers on how to effectively use telemedicine services, as it can improve their skills in self-observing symptoms and mental health management. Telehomecare can also be used to educate and promote self-care, which can lead to better health outcomes for patients. Various channels, such as phone calls, websites, apps, or chats, can be used to provide additional health care skills to caregivers [34,35].

**The Health Care Team Should Enhance Their Skills on Telehomecare Services for Patients at Home**

A technical problem has been reported as an issue. Therefore, it is crucial for the health care team to possess the necessary skills in telehomecare services to assist patients and their families when they encounter such problems in order to enhance ease of use. It has been suggested that health care teams possess the following skills when delivering telehomecare services [36]: (1) determining when to use telemedicine and assessing the ability of patients and caregivers to use it; (2) proficiency in assessing and caring remotely for patients; (3) effective communication and relationship-building with patients, caregivers, and families; (4) professionalism; (5) basic understanding of information technology; (6) knowledge of laws and privacy protection; (7) ethical considerations; (8) awareness of patient’s safety; and (9) awareness of accessibility and service culture. All of these knowledge and skills required by the health care team impact the provision of effective services which are of utmost importance. In some countries, such as the Netherlands, core competencies have been defined for nurses to ensure the effective implementation of telehealth [37].

The key findings from this study suggest that the health care team should know how to assess the situation and select the appropriate service models, whether it could be in-person home visits or telehomecare based on the experience. Although service users have different preferences, circumstances and necessities should be considered. According to the study by Doraiswamy et al [38], it was noted that telehomecare services had previously been used for noncommunicable diseases. However, during the current pandemic, they have become increasingly important for diagnosis, symptom monitoring, rehabilitation physiotherapy, and reducing the spread of pathogens as well [38]. Despite these advantages once the outbreak situation improves, telehomecare services should only be provided in appropriate cases. Patients with positive attitudes toward telehomecare services; who have used technology and have previously received in-person medical treatment; and have difficulties traveling to the hospital due to distance, finances, or health issues are typically considered suitable for telehomecare services. In addition, the content discussed during the telehomecare services should not be of a sensitive nature, particularly in the case of relationships with the caregiver [39]. It is important to ensure that the patient has a suitable place to receive telehomecare services without distractions and can provide the necessary information to the health care team.

The preferences of patients and their families are a crucial factor in choosing the right telehomecare service. It is important for the physician to consider various aspects, including the patient’s health, family preferences on service models, and the current social situation. Health care providers should engage in conversations with older adult patients and their caregivers regarding the advantages and disadvantages of telehealth, enabling patients to make informed choices between in-person and telemedicine options [23]. This is to make an informed decision and ensure patient satisfaction with the chosen service in the future.
The Barriers, Especially Those That Are Fundamental to Providing Telehomecare Services, Should Be Removed as Much as Is Feasibly Possible

In our study, we found that barriers to telehomecare can originate from patients themselves, including soft speaking voices or inadequate communication devices. Technology-related barriers, such as poor signal quality or connectivity issues, may also occur. To mitigate these issues, providers and recipients may need to invest in proper equipment, use a microphone that can absorb speech well, and place equipment in the right position to optimize the internet signal. Additionally, the financial aspect of the service should also be considered, it may be useful to have a system to restore certain devices from the service provider but must be weighed against the expense of the service provider side as well.

Some caregivers struggle to use communication devices, which can cause interruptions in telehomecare services. Patients may also feel incapable of learning to use these devices, according to a study by Huang et al. Thus, telehomecare teams should be knowledgeable about the devices used, inquire about any concerns, and provide assistance to build confidence in using them. To ensure accessibility to all areas, it is recommended to develop strong internet signal towers to support telehomecare for the national benefit.

Interestingly, this study found that the majority of caregivers for older adult patients were highly educated children, with great potential to use technology better than the patient. However, they may still have limitations in using communication devices and technology. Health care providers should not judge their ability based on age or education, and even frail older adults can be trained with help from nurses and caregivers. Each family should be evaluated on a case-by-case basis, and appropriate assistance provided.

Services Should Also Be Provided to Meet Expectations Regarding the Management of Health Problems

Caregivers of older adult patients often have specific expectations from their health care providers, such as recommendations to enhance the patient’s overall health, psychological support, and reducing transportation costs. Our findings support that caregivers perceive telehomecare as useful and prefer it when their expectations are met during the service delivery. Offering services that meet the expectations of patients and their caregivers helps improve the health care aspect and engage in service involvement, which leads to a reduction in hospitalization rates. The health care team should strive to meet these expectations to maintain the trust of patients and their families in the telehomecare health care service to the same extent as in an in-person home visit.

Strengths

The strength of this study is that the study involved caregivers with experience in both in-person home visit and telehomecare, allowing for comparison of both services. The study also focused on in-depth aspects related to caring for the bedridden older adults who may have limitations in receiving services.

Limitations

However, there are also limitations to this study. For instance, the information obtained through interviews by health care teams or personnel from the hospital may be biased. To mitigate this, the researchers took measures to ensure that the interviewer was not involved in the treatment of any of the patients and caregivers interviewed. The interview process for participants using convenience sampling may involve selecting individuals who are readily available or easy to reach, potentially leading to a sampling bias. However, the data reach its saturation. Additionally, the sex of participants was 75% (n=15) female, which may yield different insights compared with settings with varying sex ratios among caregivers. However, in Thailand, the majority of caregivers for older adults are female, with a prevalence ranging from 70% to 90%. This ratio is similar to that in our study. Given the slight tendency for male individuals to hold slightly more favorable attitudes toward technology use than female individuals, although not substantially, it is important to note that this sex imbalance among caregivers should not significantly impact the study's results. The participants of this study were mainly caregivers of older adults with stable chronic illnesses. Caregivers of patients with more complicated illnesses or the terminally ill may have different attitudes or expectations. We did not include the attitudes of the medical service team with regard to this aspect. These attitudes may affect the outcomes of services and treatment as well, which may be subject to further study. Last, conducting a study comparing telehomecare with no care would underscore the significance of home-delivered care. However, since our initial aims did not include a comparison of these 2 populations, we did not collect data from individuals who have not received the service. It may be worthwhile to consider further research involving families of older adults with health conditions that limit their function but are unable to obtain home care services. Nevertheless, obtaining information about this population could be challenging.

Conclusion

In conclusion, telehomecare is a viable option to supplement in-person home visits. It has the potential to provide quality health care services to patients and allow health care teams to offer treatment and advice remotely. Caregivers have shown a positive attitude toward telehomecare, which is comparable in effectiveness to in-person visits. Despite potential challenges, telehomecare can be a useful alternative in situations where in-person visits are not possible. However, it is important to enhance the skills of caregivers and health care teams to overcome barriers and optimize the use of telehomecare.
Acknowledgments
This work was partially supported by Chiang Mai University.

Data Availability
The data sets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Framework matrix.

References


https://aging.jmir.org/2024/1/e48132
Abbreviations

SRQR: Standards for Reporting Qualitative Research

©Pansiree Onseng, Wichuda Jiraporncharoen, Sasiwimon Moonkayaow, Pimchai Veerasirikul, Nutchar Wiwatkanupakarn, Chaisiri Angkurawaranon, Kanokporn Pinyopornpanish. Originally published in JMIR Aging (https://aging.jmir.org), 07.02.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
A Web-Based Intervention Based on Acceptance and Commitment Therapy for Family Caregivers of People With Dementia: Mixed Methods Feasibility Study

Golnaz. L Atefi1, BSc, MSc; Rosalia J M van Knippenberg1, BSc, MSc, PhD; Sara Laureen Bartels1, BSc, MSc, PhD; Andrés Losada-Baltar2, BSc, MSc, PhD; María Márquez-González3, BSc, MSc, PhD; Frans R J Verhey1, BSc, MSc, PhD, MD; Marjolein E de Vugt1, BSc, MSc, PhD

1Department of Psychiatry and Neuropsychology, Alzheimer Centrum Limburg, Maastricht University, Maastricht, Netherlands
2Departamento de Psicología, Facultad de Ciencias de la Salud, Universidad Rey Juan Carlos, Madrid, Spain
3Biological and Health Psychology Department, School of Psychology, Universidad Autónoma de Madrid, Madrid, Spain

Corresponding Author:
Golnaz. L Atefi, BSc, MSc
Department of Psychiatry and Neuropsychology
Alzheimer Centrum Limburg
Maastricht University
Dr. Tanslaan 12
Maastricht, 6229 ET
Netherlands
Phone: 31 (0) 43 388 1022
Email: g.atefi@maastrichtuniversity.nl

Abstract

Background: Acceptance and commitment therapy (ACT), as an empirically based third-wave cognitive behavioral therapy, has shown promise in enhancing well-being and functioning across diverse populations. However, in the context of caregiving, the effect size of available ACT interventions remains at best moderate, sometimes accompanied by high dropout rates, highlighting the need for more effective and feasible intervention designs.

Objective: The objective of our study was to evaluate the feasibility and acceptability of a fully online ACT program designed for family caregivers of people with dementia. This study aimed to boost psychological flexibility and support caregivers, enabling them to realize and prioritize their own life values alongside their caregiving responsibilities.

Methods: A mixed methods feasibility study using an uncontrolled pretest-posttest design was conducted. This intervention included a 9-week web-based self-help program based on ACT incorporating collaborative goal setting and weekly web-based motivational coaching for family caregivers of people with dementia. This study involved 30 informal caregivers recruited through memory clinics and social media platforms in the Netherlands and received approval from the Medical Ethics Committee of the Maastricht University Medical Center+ (NL77389.068.21/metc21-029).

Results: A total of 24 caregivers completed the postintervention assessment, indicating a high adherence rate (24/29, 83%). Caregivers reported positive feedback regarding collaborative goal setting, but some found challenges in implementing new skills due to their own habitual responses or the unpredictable context of dementia caregiving. Personalizing the intervention based on individual value preferences was highlighted as beneficial.

Conclusions: Compared to other web-based self-help ACT interventions for family caregivers, this intervention showed a high adherence and sufficient level of feasibility, which underscores the use of personalization in delivering web-based interventions. Moreover, the potential of this ACT-based intervention for family caregivers of people with dementia was demonstrated, suggesting that further research and a larger-scale controlled trial are warranted to validate its effectiveness.

International Registered Report Identifier (IRRID): RR2-10.1136/bmjopen-2022-070499

(JMIR Aging 2024;7:e53489) doi:10.2196/53489

https://aging.jmir.org/2024/1/e53489
KEYWORDS
acceptance and commitment therapy; ACT; psychological flexibility; behavior change; theory-guided eHealth; web-based intervention; supported self-help; family caregivers; dementia

Introduction

Background

The number of people with dementia is predicted to double every 20 years, which will lead to a corresponding rapid increase in the number of family caregivers [1]. Family care increases the quality of life of people with dementia and reduces formal care costs, thus making a substantial contribution to dementia care management. This is noteworthy as most people with dementia rely on a range of crucial and unpaid support from family caregivers [2]. Although the experience of caregiving might be fulfilling and positive, research also shows that, with advancing dementia, the need for care can become increasingly time-intensive, stressful, and more complex, posing a risk to the overall well-being of caregivers. In the context of dementia, family caregivers may experience inevitable and long-lasting changes due to the progressive nature of the condition and the increasing dependence of the person with dementia on their caregivers.

Importantly, the adverse effects of caregiving can be addressed and improved through a wide range of psychological interventions [3-6]. However, the long-term care situation and associated (inevitable) changes in caregivers’ lives underscore the importance of acceptance-based interventions that focus on developing skills to effectively address the management of maladaptive thoughts and emotions and acceptance of ongoing changes [7]. Specifically, the ways through which caregivers perceive and respond to internal (ie, thoughts and feelings) and external (ie, environmental) stressors are significant predictors of negative outcomes in this population [8,9]. Subsequently, acceptance and commitment therapy (ACT), as an empirically based third-wave cognitive behavioral therapy, might be particularly noteworthy due to its scalable focus on promoting psychological flexibility [10]. From the ACT perspective, psychological flexibility refers to efficient functioning in the presence of difficult experiences and is achieved via 6 interrelated core processes. These processes include openness to internal experiences (ie, acceptance), defusing from thoughts (ie, cognitive defusion), being in the present moment and aware of oneself and others (ie, mindfulness), having a distinct perspective on internal experiences (ie, self as context), identifying meaningful action qualities that can be connected with bringing purpose and motivation (ie, values), and active engagement and behavioral action aligned with values (ie, committed action) [11]. Hundreds of randomized controlled trials (RCTs) have demonstrated the effectiveness of ACT in improving overall well-being in the general population and in people with mental or somatic health problems [6,12].

ACT for Family Caregivers of People With Dementia

In the context of dementia caregiving, ACT shows promise in promoting acceptance of change and increasing willingness to take meaningful actions, leading to improved psychological flexibility and better overall functioning in caregivers [13-15].

Several RCTs have demonstrated that face-to-face individual ACT interventions provided by trained therapists result in a significant reduction in depression and anxiety in family caregivers of people with dementia compared to control groups [13,15,16]. Furthermore, non-RCT ACT studies have also shown promise in supporting family caregivers of people with dementia through modalities such as telephones [17], videoconferencing [18], group settings [19], or web-based self-help modules [20,21].

eHealth Adaptations for Family Caregivers

In recent years, technological advancements have facilitated the development and adaptation of a wide range of acceptable and promising psychological programs, from face-to-face to eHealth (ie, the use of internet to promote well-being) [4,22,23]. In particular, web-based self-help interventions are cost-effective and accessible approaches that reduce the significant involvement of health care professionals, allowing caregivers to complete the intervention on their own time using computers, tablets, or mobile devices. Web-based self-help interventions provide materials such as modules, text, and videos, enabling users to navigate the program at their own pace [24].

However, the effect sizes of the available interventions are still at best moderate [3] with a high dropout rate [20], indicating the need for more effective intervention designs [15]. Furthermore, in the context of caregiving, family caregivers of people with dementia are often older adults. The potential lower digital literacy and preference for traditional face-to-face psychological support, coupled with implementation limitations, make them the subgroup of caregivers who most frequently report challenges related to accessing and using eHealth interventions [4]. Thus, in the context of eHealth, guided self-help interventions including “minimal contact” might be a promising approach for this population. In a “minimal contact” approach, health care professionals (eg, coaches) are involved for nontherapeutic purposes and mainly for periodic check-ins, teaching participants how to use the digital tools, and provision of initial rationales [23,25].

Personalizing Self-Help Interventions From the ACT Perspective

Personalizing interventions is an important and effective strategy to increase adherence and prevent dropout rates, making trials, including RCTs, more feasible, acceptable, and effective [26]. One effective and efficient approach to personalization is collaborative goal setting, in which individuals, together with health care professionals, set specific goals based on their own needs and resources before the intervention, leading to increased motivation and, ultimately, behavior change [27]. From the ACT perspective, there is a distinction between goals and value-based choices and their impact on (long-term) behavior change. Goals are typically external and have a clear end point that can be achieved or completed. In contrast, values are meaningful qualities that cannot be obtained or finished but rather help set meaningful goals and guide long-term patterns.
of behavior. Thus, acknowledging personal values and nesting specific goals underneath them is more likely to drive effective behavior change in the long term [11]. Engaging in value-based activities in the context of caregiving has a positive association with emotional well-being [28] and a negative association with distress [16].

To our knowledge, collaborative goal setting based on caregivers’ values has not yet been explored within the context of web-based ACT self-help interventions. Gaining further insights into the feasibility of web-based ACT interventions for family caregivers of people with dementia might contribute to decisions related to the implementation of ACT interventions in clinical practice, facilitating intervention refinements and, ultimately, leading to the design and development of more acceptable, effective, and sustainable interventions for future large-scale controlled trials.

This Study
Despite the growing literature on ACT and the efficacy of eHealth interventions in family caregivers, there is a limited focus on understanding the specific challenges and opportunities of web-based self-help interventions for this population [29]. To date, few studies have used supported or personalized self-help ACT to address the specific needs of family caregivers. The aforementioned studies, although promising [21], showed a high dropout rate [20], or the intervention did not prove highly effective [30]. Therefore, in response to the need for a more efficient intervention design for family caregivers of people with dementia, this pilot trial aimed to use a mixed methods assessment approach (qualitative and quantitative data) to examine the feasibility and acceptability of the ACT for informal caregivers of people with dementia intervention. This web-based guided intervention is designed for community-based family caregivers of people with dementia through web-based ACT modules, collaborative goal setting based on individuals’ personal values before the intervention, and minimal-contact motivational coaching during the intervention. Subsequently, this study aimed to address the following research question: what are the practical and conceptual barriers and facilitators influencing the feasibility and acceptability of this guided web-based intervention?

Methods

Overview

This study was a pilot trial with a mixed methods approach, a baseline assessment, a 9-week web-based intervention embedded with web-based weekly motivational coaching, and a postintervention assessment. This study was reported according to the guidelines presented in the CONSORT (Consolidated Standards of Reporting Trials) checklist [31]. The CONSORT checklist is presented in Multimedia Appendix 1 [32]. The complete design is described in the protocol [33], and the methodological details relevant to this feasibility and acceptability study are presented in the following sections.

Participants

A sample size of 30 participants is considered to be sufficient to enable a reasonable calculation of the key factors relevant to feasibility (eg, attrition rates) and provide useful information required for recommending a larger controlled trial [34,35]. Therefore, 30 family caregivers of people with dementia were recruited for this study.

Eligible family caregivers were recruited sequentially from May 2022 to June 2023. The following inclusion criteria were applied: (1) adult caregivers (aged ≥18 y), (2) self-identified primary family caregivers of a person with a diagnosis of dementia, (3) caring for the care recipient for at least 3 hours per week for at least 3 months, (4) internet and tablet or computer accessibility in the household, and (5) consent to participate. Family caregivers were excluded if (1) they indicated cognitive difficulties or disorders in their medical record (based on self-report) or (2) they had undergone psychotherapy or psychopharmacological treatment during the previous 3 months.

Recruitment Procedure and Screening

Family caregivers were recruited from referrals by clinicians (eg, psychiatrists or psychologists) at the memory clinic of the Maastricht University Medical Center+ in the Netherlands. Furthermore, recruitment took place using printed or web-based flyers and website posts by patient and caregiver support organizations in the Netherlands, the Dutch Alzheimer Association, and local mental health institutions. Information about the study as well as a self-addressed stamped envelope for returning the informed consent form were provided to all the participants via post. When the research team received the signed informed consent form (in paper format), the study officially started. The process of screening and recruitment is shown in Figure 1, and further details on the recruitment procedure can be found in the study protocol [33].
Intervention

Overview
The guided web-based intervention followed a written protocol [33]. As a general structure, each week, participants first completed a web-based module and then received a telephone or video call from a motivational coach. Participants were instructed to complete consecutive steps, including (1) a 50-minute web-based video call at baseline with a motivational coach for value-based goal setting, (2) 9 self-help web-based ACT modules (20-30 min each), and (3) an approximately 20-minute weekly web-based video call with a motivational coach for 9 weeks. In addition, participants had the flexibility to extend the duration of the intervention and complete the 9 modules and coaching sessions in 12 weeks. Contact with the coach took place in individual sessions.

Collaborative Goal Setting
Collaborative goal setting [36] was chosen to personalize the guided web-based intervention. During the individual value-based goal-setting session with the motivational coach, each participant set one value-based personalized goal that was specific, measurable, attainable, realistic, and time-orientated (SMART) [37]. Individuals could select from a predefined list of value-based actions or write goals in their own words (Multimedia Appendix 2). The list was prepared based on the ACT model and an adapted version of the valued living questionnaire for dementia caregiving [28,29,38]. Specific value-based actions with clear steps were defined during the collaborative goal-setting session and were embedded into future activities within the intervention. The steps of collaborative goal setting are shown in Multimedia Appendix 3.

Web-Based Self-Help Modules
Details about the intervention can be found elsewhere [33]; in brief, caregivers focused on 1 of the 6 ACT components (ie, acceptance, cognitive defusion, mindful attention, self as context, value, or committed action) in each module, gradually covering all the core principles of psychological flexibility. The 9 self-help web-based modules were released weekly consisting of a standard structure including a brief introductory text, a short video, an ACT-related metaphor, and content-oriented assignments. Module 1 provided a short introduction to ACT and a program overview. In module 2, “Creative Hopelessness” was introduced to reflect on the dysfunctionality of avoidance strategies for difficult thoughts and feelings in life. Modules 3 to 8 reviewed each of the 6 ACT components in detail. Finally, session 9 was a recap of all ACT components incorporating strategies for relapse prevention in everyday life (Multimedia Appendix 4). In addition, ACT metaphors were embedded in the content of the modules. For example, the boomerang metaphor was used for avoiding unpleasant feelings. The concept is that the more one attempts to throw the metaphorical boomerang away, the more forcefully it will eventually return [39].

After completing a module, caregivers received an automated email notification from the website reminding them of the availability of the next module. Access to the modules was sequential, requiring completion in a specific order rather than allowing access to all modules at once. Further details on the content of the website are presented in Table 1.
Table 1. Overview of the 9 modules included in the intervention.

<table>
<thead>
<tr>
<th>Module</th>
<th>Title</th>
<th>ACT³ strategy</th>
<th>Example of the exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Introduction</td>
<td>The introduction of ACT and the content of the program consists of modules with videos explaining ACT components accompanied by metaphorical images and exercises to enhance personal resilience.</td>
<td>“You are now invited to answer these two questions for yourself: Where do you want to go with your life (as a person, privately and in your work)? What stands in your way?”</td>
</tr>
<tr>
<td>2</td>
<td>Creative hopelessness</td>
<td>Exploring creative hopelessness strategies among informal caregivers for navigating unpleasant thoughts and feelings.</td>
<td>“Share three personal examples of how you prevent unpleasant feelings, employing strategies like prevention, avoidance, and reduction. This exercise fosters self-awareness about your functioning in challenging situations.”</td>
</tr>
<tr>
<td>3</td>
<td>Acceptance</td>
<td>Recognizing caregivers’ potential struggles with negative emotions, this module involves fostering acceptance by making room for unpleasant feelings.</td>
<td>“This exercise encourages reflection on personal pain, strategies for coping, and resulting suffering. Identify specific instances of pain, whether thoughts, feelings, physical sensations, or situations.”</td>
</tr>
<tr>
<td>4</td>
<td>Defusion</td>
<td>Addressing the tendency of individuals to take their thoughts seriously, the focus is on defusion—encouraging individuals to distance themselves from difficult thoughts.</td>
<td>“Write a negative self-view, e.g., ‘I’m not good enough,’ then rephrase it as ‘I have the thought that I am not good enough.’ Notice the difference between directly identifying with the thought and acknowledging it as a passing idea.”</td>
</tr>
<tr>
<td>5</td>
<td>Self as context</td>
<td>Acknowledging the tendency of individuals to define their self-image based on perceived expectations, the emphasis is on the self as context—creating room for individuals to be themselves and fostering flexibility in their self-image to alleviate stress.</td>
<td>“In ACT, your self-image is like wearing a tailor-made suit that may not fit your current feelings. Reflect on times this felt restrictive and consider less uncomfortable suits you’ve worn. Explore a more flexible approach and think of the benefits of taking them off for a moment.”</td>
</tr>
<tr>
<td>6</td>
<td>Mindfulness</td>
<td>Recognizing that excessive focus on the past or future may not always be helpful, the emphasis is on the here and now—encouraging individuals to pay sufficient attention to the present moment, the only time when we can truly live, act, and experience.</td>
<td>“Reflecting on past and future often overshadows the present; embracing the moment enhances life satisfaction. Losing the present in thoughts might result in missing the richness of the present experience. Try to recognize an example of this in your own life.”</td>
</tr>
<tr>
<td>7</td>
<td>Value</td>
<td>Recognizing the significance of acknowledging what truly matters in one’s life, the focus is on values—actively encouraging individuals to ask whether their values are sufficiently present in their lives and put them into practice.</td>
<td>Value-sorting task: “Identify core life values using a set of cards, categorizing them as ‘Very Important,’ ‘Important,’ or ‘Not Important,’ and then prioritizing the top 10 values to actively invest in.”</td>
</tr>
<tr>
<td>8</td>
<td>Committed action</td>
<td>Emphasizing the facilitation of a more meaningful life, the focus is on committed action—encouraging individuals to define concrete and feasible actions toward their values, actively invest in them, and translate them into value-based actions.</td>
<td>“Take practical steps and reflect on each significant value and ask, ‘What is the smallest, easiest step I can take in the next 24 hours to align my life with that value?’ Ensure specificity and feasibility in these actions for practical implementation.”</td>
</tr>
<tr>
<td>9</td>
<td>Psychological flexibility</td>
<td>The concluding module highlights that cultivating psychological flexibility through the practice of the 6 core skills enables individuals to address problems differently and with greater flexibility, enriching their lives in ways that hold personal value.</td>
<td>“You were introduced to all the different ACT components. We invite you to answer the initial questions from Module One again: Where do you want to go with your life (as a person, privately and in your work)? What stands in your way?”</td>
</tr>
</tbody>
</table>

³ACT: acceptance and commitment therapy.

Weekly Coaching
A trained research assistant with experience in intervention coaching was appointed from Maastricht University to fulfill the role of the motivational coach. As part of the weekly coaching process, participants were encouraged to complete the weekly module before scheduling a video call. During the weekly coaching, the coach inquired about the participants’ overall experience with the web interface (ie, whether any technical or other issues occurred), how goal attainment was progressing (ie, value-based actions defined during the collaborative goal setting), and whether they experienced a (positive or negative) change in their motivation (Multimedia Appendix 5). Weekly coaching was not intended as a therapeutic function but rather to provide technical support and motivate participants to follow the web-based self-help modules (ie, adherence) and stay engaged with their value-based goals. Thus, goal attainment and module completion were not mandatory before coaching took place.

Sampling and Intervention Quality
Data were collected using the Castor web-based platform provided by Maastricht University. Intervention integrity was monitored independently by experts. Several types of monitoring...
visits were conducted by the Clinical Trial Center Maastricht for the purpose of quality and control at the beginning of the study (ie, site initiation visit); during the course of the study (ie, interim monitoring visits); and at the end of the study, when participants had completed the study and all data had been collected (close-out visit) [33]. Quantitative and qualitative data were used to inform the feasibility and acceptability of the intervention for the participants and coach. The number of referrals, number of eligible participants willing and not willing to take part, reasons for declined participation or dropout after signing the informed consent form and before starting the intervention (if provided), and amount of time spent on recruiting 30 participants were monitored during recruitment. Quantitative and qualitative data on the intervention procedure were collected to inform use logs and interaction with the web-based modules. Barriers (eg, technical problems or difficulties) to user engagement were qualitatively collected during weekly coaching. To distinguish between intervention rejections and dropouts, we defined dropout as attrition following the start of the intervention (ie, those participants who attended at least one session of the intervention but discontinued it) [26]. Reasons for dropout after starting the intervention (if provided), attrition rate of weekly coaching sessions, and ACT module completion were also recorded.

Assessment

Overview

This study specifically focused on feasibility and acceptability aspects. The preliminary efficacy and changes in psychological outcomes will be reported elsewhere. Further details on the measures of preliminary efficacy are described elsewhere [33].

End-of-Module Questionnaire

Following the previous study, the research team developed the end-of-module questionnaires [40] with the goal of gaining additional insights into the feasibility and acceptability of each module. Subsequently, content comprehension was assessed directly after completion of each module via a feedback questionnaire. This questionnaire included 3 items (ie, “I found today’s module useful,” “I have experienced the content of the modules as stressful,” and “I can apply the content of today’s modules in my daily life”) rated on a Likert scale ranging from 1=strongly disagree to 7=strongly agree on perceived usefulness, stressfulness, and applicability of each module to one’s daily life.

Goal Attainment

The level of goal achievement, as well as qualitative feedback on the feasibility and acceptability of goal achievement, was collected on a weekly basis during coaching. Goal attainment scaling was mapped on a prespecified ordinal scale, and the number of attainment levels ranged from −3 to +2. Each SMART goal was set at different layers as an “action list” including different levels of goal attainment. Level “0” was set as the “expected” level, and the rest of the levels were defined by a possible change in goal attainment. Any progress from the “expected” level was scored with “+1” as the “better than expected” level or “+2” as the “much better than expected” level. Deterioration in goal attainment was scored with “−1” as the “improvement but less than expected” level. The “−2” score was assigned to the “current” level and addressed “no change” from the goal-setting day, and “−3” referred to the “much less than expected” level in goal attainment. Setting an in-between “−2” score as the “current” level was considered to prevent floor effect and capture deterioration from the “current” individuals’ state [37]. During collaborative goal setting at baseline, each caregiver defined and clarified 5 layers of their SMART goals (from −3 to +2).

The Program Participation Questionnaire

Further insights into the usability, clarity, and acceptability of the intervention for family caregivers was obtained during a postintervention semi-structured interview using the Program Participation Questionnaire (PPQ) [40]. The PPQ consisted of 26 items scored on a Likert scale ranging from 1=strongly disagree to 7=strongly agree focusing on 3 main areas, including the applicability of the intervention in everyday life, feasibility, usability, acceptability, and content quality and quantity (Multimedia Appendix 6). Furthermore, the number of log-ins to the modules and feature use were collected and compared with self-reported data. The feasibility and perceived experience of the coach was evaluated using a brief 6-item coach questionnaire focusing on the intervention’s usability and relevance for the coach, general perceived experience, the program’s positive and negative aspects, and suggestions for improvements (Multimedia Appendix 7).

Demographic Information

During the baseline assessment, family caregivers completed a demographics questionnaire providing information about their age, sex, level of education, living situation (ie, whether the caregiver and the person with dementia lived together or independently), hours of caregiving per week, type of kin relationship with the person with dementia, years since diagnosis, and dementia type.

Data Analysis

The PPQ was analyzed quantitatively and qualitatively. Due to the lack of external criteria to properly define feasibility [41], median scores as a conventional strategy were defined as determinants of the overall feasibility, usability, and acceptability [40,42]. This approach to evaluating feasibility was also previously used in a Delphi study [43]. The overall PPQ scores ranged from 26 to 182. The median score of 130 was deemed as the cutoff and, thus, “acceptable feasibility” [40]. Mean item scores (range 1-7) of <5 (“slightly agree”) were considered as having potential for improvement. Participants further reflected qualitatively on their scores, and their reflections were audio recorded and transcribed verbatim. The deductive content analysis was conducted by authors GLA and RVK using field notes to interpret the quantitative scores on the PPQ [42,44]. In addition, to evaluate whether the intervention components aligned with the specific needs of the target group [45], field notes were classified and deductively coded into four main categories to address the intervention components: (1) overall experience and suggestions, (2) goal
setting and value identification, (3) web-based self-help ACT modules, and (4) coaching. Furthermore, the self-report acceptability questionnaires at the end of each module were summarized in SPSS (IBM Corp) using descriptive statistics. Finally, web feature use and the number of log-ins were collected and subsequently compared with self-reported data. Data from the weekly coaching sessions were incorporated to complement the log data and the postintervention semistructured interviews.

Ethical Considerations
The study was approved by the Medical Ethics Committee of the Maastricht University Medical Center+ (NL77389.068.21/metc21-029). All participants provided informed consent, and the rigorous protection of privacy and confidentiality of participants was safeguarded throughout the study. Data were anonymized when applicable to safeguard participant identities. As a token of appreciation, participants received a €25 (US $27.13) gift card for their involvement.

Results
Overview
A total of 33 family caregivers were deemed eligible to participate in the study, of whom 30 (91%) provided informed consent and 24 (73%) completed the postintervention assessment (Figure 2).

In total, 3 eligible caregivers declined to sign the informed consent form citing the extra burden on their caregiving responsibilities (n=2, 67%) and the admission of the person with dementia to a nursing home (n=1, 33%) as reasons for their decision. Of the 29 participants who initiated the intervention, 4 (14%) decided to leave prematurely. Thus, based on the definition of dropout in this study, an adherence rate of 83% (24/29) was recorded, including 1 withdrawal and 4 dropouts. Of the 30 caregivers who signed the informed consent form, 1 (3%) did not continue the baseline assessment due to difficulties in arranging an electronic device, such as a laptop or tablet, to continue the study. In addition, of the remaining 29 caregivers, 1 (3%) withdrew due to the death of the person with dementia, and 4 (14%) dropped out citing concerns that the study was too time-consuming or emotionally challenging or added an extra burden to their existing caregiving responsibilities. Examples of reasons for dropout included the following:

I wasn’t that fond of it. It makes you depressed, all that thinking about the past and stuff. It’s not what I expected. [ACT-IC015; aged 63 years; male]

I don’t want to dwell on it, don’t want to learn to think differently either, I find that scary. [ACT-IC018; aged 51 years; female]

No differences were identified between caregivers who dropped out and those who completed the study. Of the 24 participants who completed the study, most were female (18/24, 75%), and the hours of caregiving were often reported to be >15 per week, ranging from 3 to 20. The duration of dementia (ie, time since diagnosis) was, on average, 4.1 (SD 2.7) years, with a higher prevalence of Alzheimer (12/24, 50%) followed by vascular dementia (6/24, 25%). Further sociodemographic characteristics of the caregivers are shown in Table 2.
Figure 2. Participant flow.

- Recruited from social media (n=22)
- Recruitment from clinicians (n=11)

- Eligible to participate (n=33)

- Consented to participate (n=30)
  - Lost to baseline assessment (n=1)
    - Reasons for dropout: Difficulty in arranging laptop or tablet
  - Withdrawn from the study due to the death of the person with dementia (n=1)
  - Dropout (n=4)
    - Reasons for dropout: Too confronting, burdensome or time-consuming

- Completed baseline assessment (n=29)

- Completed post-intervention assessment (n=24)
Table 2. Overview of the sample characteristics (N=24).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y), mean (SD)</td>
<td>62.5 (13.1)</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>18 (75)</td>
</tr>
<tr>
<td>Male</td>
<td>6 (25)</td>
</tr>
<tr>
<td><strong>Kin relationship, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Spouse</td>
<td>14 (58)</td>
</tr>
<tr>
<td>Child</td>
<td>9 (38)</td>
</tr>
<tr>
<td><strong>Ethnicity, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>24 (100)</td>
</tr>
<tr>
<td><strong>Education, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Completed college</td>
<td>15 (62)</td>
</tr>
<tr>
<td><strong>Employment status, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Currently employed</td>
<td>10 (42)</td>
</tr>
<tr>
<td><strong>Living arrangement, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Living with care recipient</td>
<td>17 (71)</td>
</tr>
<tr>
<td><strong>Hours of caregiving per week, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>≥15</td>
<td>15 (62)</td>
</tr>
<tr>
<td><strong>Dementia diagnosis, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Alzheimer disease</td>
<td>12 (50)</td>
</tr>
<tr>
<td>Vascular</td>
<td>6 (25)</td>
</tr>
<tr>
<td>Young onset dementia</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Frontotemporal</td>
<td>1 (4)</td>
</tr>
<tr>
<td>Parkinson disease</td>
<td>2 (8)</td>
</tr>
<tr>
<td>General (diagnosed, not specified)</td>
<td>2 (8)</td>
</tr>
<tr>
<td><strong>Duration of dementia (y), mean (SD)</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.1 (2.7)</td>
</tr>
</tbody>
</table>

Results of PPQ

*Overview*

The total score on the PPQ ranged from 90 to 182. The average total score on the PPQ of the 24 caregivers who completed the intervention was 163.4 (SD 22.3). On average, family caregivers gave all items a score of at least 5 (mean 6.29, SD 0.46), with a total median score of 172, indicating high perceived feasibility and acceptability (Figure 3).
The quantitative results of the PPQ on intervention feasibility and acceptability were consistent with postmodule self-report acceptability as well as with deductive qualitative content analysis of the field notes.

Qualitative feedback informed the feasibility and acceptability of the intervention components, specifically (1) overall experience, (2) goal setting and value identification, (3) web-based self-help ACT modules, and (4) coaching. In addition, a semistructured interview was conducted with the motivational coach to gain more insights into study feasibility for the interventionist.

**Overall Experience**

Overall, participants found the intervention’s design comprehensive, logically structured, and strongly interconnected, “like a string of chain,” emphasizing that removing a single component would disrupt the functionality of the intervention. Enhanced adaptability to the caregiving role and improved ability to focus on meaningful life values while providing care for people with dementia were noted as valuable components in addressing the balance between caregiving and self-care. Caregivers reported that their enhanced psychological flexibility, achieved through ACT principles, supported their resilience and also positively influenced people with dementia, describing it as a “win-win” situation:

> It was aligned, enriching, confronting, and educational. Without a coach, it would have been easier to let things slip. Providing care becomes much more sincere and high quality by being mindful. You can only do that if you have space for it, so you have to let go of your need for control. It starts with yourself, and you can apply it everywhere, that is a relaxing feeling. I learned to deal with myself in a different way, I could respond more calmly. I wasn’t overwhelmed and overstimulated; it had such a positive effect on the person with dementia. [ACT-IC 002; aged 47 years; male]

Caregivers suggested the inclusion of a psychoeducational module specifically addressing the application of ACT to the unique challenges caregivers face in dementia care. Such a module may support contextualizing ACT and enhance the understanding of the metaphors used within the intervention, making it more relevant and applicable to the unpredictable and progressive nature of the condition. In addition, some caregivers expressed the desire to have ongoing access to the modules or receive a printed booklet containing all the modules, allowing them to reference the material even after the study. While acknowledging that the intervention was confronting and required more time than initially anticipated, participants still expressed their recommendation of the intervention to other caregivers.

**Goal Setting and Value Identification**

Family caregivers expressed that setting goals before participating in the intervention was essential for their engagement with the program and establishing a road map toward desired behavior change. Formulating SMART goals made the process more tangible for participants and stimulated a sense of purpose during the intervention, which led individuals to plan and adapt self-management and personal values to their caregiving responsibilities. Most caregivers often emphasized...
the importance of receiving guidance in value identification, prioritizing goals, breaking value-based actions down into manageable steps, and adjusting goals when necessary:

Providing guidance is important because then I receive a kind of confirmation that I should do it as well. Otherwise, it remains so open-ended. The coach allowed me the freedom to set my own goals and provided guidance when needed. Setting goals and timeframes helped me avoid feeling guilty about taking time for myself. [ACT-IC012; aged 65 years; female]

Goal Attainment

Participants had mixed experiences when it came to engaging with the action list, which included layers of goal attainment, as well as when they were asked to self-report their progress on their goals during meetings with the coach. Some caregivers found the action list to be beneficial for adapting their goals to better align with the changing demands of caregiving, enhancing their ability to navigate through personal values and caregiving responsibilities effectively. Nonetheless, some caregivers encountered challenges in both accomplishing their goals and accurately assessing their level of achievement. Furthermore, a desire to modify their goals was also noted among some participants. Factors that influenced the attainment of goals primarily included caregiving demands such as the inability to leave the care recipient alone, their own health conditions (eg, physical pain), and time limitations. These factors often resulted in adjustments to the goals throughout the intervention. In addition, a shift in focus to a different set of values was another factor that influenced goal adjustment.

Reduced engagement with the action list or failure to attain goals at a desirable level for some caregivers caused feelings of anxiety or increased burden when reporting to the coach:

I had the tendency to give socially desirable answers when the coach asked if I had worked on my goals. I would be asked if I had gone cycling, but I can’t leave my partner alone. So, I do want to take more time for myself, but the situation doesn’t always allow me to engage in outdoor activities. [ACT-IC 009; aged 74 years; female]

According to the coach, using the SMART framework for goal setting was valuable for structuring collaborative goal-setting sessions. However, adjusting goals to account for comorbidities and caregiving demands presented challenges at times in providing guidance, particularly within limited time frames.

Web-Based Self-Help ACT Modules

Content

The content of the web-based self-help ACT modules was reported mainly as “easy to follow” and helped caregivers in adapting their coping strategies, acquiring new skills to navigate challenges, cultivating mindfulness of the present moment, and reframing their perspective to align with their life values. According to caregivers, following ACT principles in parallel with goals empowered them to gain more insights into psychological flexibility, enabling them to take practical steps toward living in accordance with their values:

I went out of my comfort zone...I usually think that if you want something, you’ll just do it, but it (i.e., ACT) was truly an eye-opener. I was constantly realising I need to do this, oh yes, I need to pay attention to that as well, of course. Those moments of realisation, I found them very valuable. [ACT-IC 004; aged 52 years; female]

After completing each module, participants provided self-report feedback and rated the usefulness, stressfulness, and applicability of the content on a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree). All participants (24/24, 100%) completed all modules along with the self-report questionnaires. According to their feedback, the web-based self-help modules were generally perceived as useful, with a mean score of 5.6 (SD 0.2; range 5.4-6.1), and applicable to daily life, with a mean score of 5.2 (SD 0.5; range 4.4-5.9). The stressfulness of the content was rated with a mean score of 3.5 (SD 0.8; range 2.4-4.4). Among the various components, the “self as a context” component was identified as the most useful yet one of the most stressful modules after “Acceptance” (Multimedia Appendix 8).

Overall, participants expressed that the ACT modules enabled them to “change attitude,” which was necessary for openness and developing skills that are “potentially applicable” in everyday life. However, taking perspective did not always result in actual changes in behavior in everyday life for some (older) caregivers. The most frequently reported barrier to applying new skills “in the situation” was the role of habits. Caregivers often highlighted that implementing new ACT skills took time to “sink in,” particularly due to habitized responses to specific cues that developed through years of repetition:

You’ve been doing things your whole life, often unconsciously, and now you’re becoming aware of things. Becoming aware and then changing your behavior takes time. [ACT-IC001; aged 65 years; female]

In difficult situations, it’s not so easy to implement everything you know. You know it now, but it can still be challenging to apply in the situations. [ACT-IC006; aged 71 years; female]

With a certain age and lifestyle, it’s a significant adjustment, not easy, but the adaptations are necessary. [ACT-IC011; aged 79 years; male]

According to the coach, extending the completion time frame to 12 weeks allowed caregivers more time and space to learn the materials. However, given the time required to acquire new skills, extending the time frame to >12 weeks could also be beneficial.

Format

Most participants found 1 module per week to be suitable, with some suggesting that 2-week intervals could also work. In total, 29% (7/24) of the caregivers required >9 weeks to complete all 9 modules. Caregivers found the web-based format beneficial as it eliminated the need for additional travel time, making it...
convenient to integrate the intervention into their daily routines. In addition, the convenience of the self-help modules was appreciated, which allowed caregivers to review the material and take notes at their own pace, offering an opportunity to consider personal preferences:

I have difficulty remembering certain things (concerning the intervention material). I have to reread things to apply them properly. [ACT-IC001; aged 65 years; female]

All participants (24/24, 100%) completed the intervention using computers or laptops with internet access. Regarding the experience of the web-based format, caregivers expressed that the presence of a coach supported them with navigating through the web setup. The combination of video, text, and assignments was expressed as useful to comprehend ACT exercises and metaphors. However, some caregivers chose to review the material either by accessing the web application on their mobile phones or offline by printing the material rather than logging into their accounts via their computers:

I printed it out so I could read it and write down my thoughts and notes. I prefer paper over online, which might have to do with my age. [ACT-IC006; aged 71 years; female]

According to the coach, addressing technical difficulties provided the necessary support for participants with lower digital literacy and prevented dropout due to technical issues.

**Coaching**

Overall, participants found the coaching sessions aligned, enriching, and constructive. The supportive listening, availability of the coach through multiple channels (video call, phone call, and email) during coaching, technical support for the web-based module, feeling safe with privacy on the web, and flexible rather than fixed coaching appointments were particularly appreciated. Caregivers reported that receiving personal feedback enhanced their active engagement with the entire intervention. This engagement, in turn, increased their motivation to learn new perspectives presented by the ACT modules and strengthened their commitment to achieving their goals. The contact with the coach through digital means was mentioned as added “accountability” and a suitable modality, particularly for those who preferred direct contact with health care professionals.

**From the Coach’s Perspective**

The weekly coaching sessions revealed several prominent barriers faced by caregivers, including technological, intrinsic, and extrinsic challenges. Technological barriers centered on issues related to digital literacy, limited access to digital devices, and a lack of self-efficacy in technology use. The coach addressed these concerns by offering supplementary guidance, providing the option to print materials for offline use, and enabling access to the intervention through a mobile-friendly web application. Intrinsic obstacles included motivational struggles, physical discomfort, and reduced sensory abilities, which influenced the caregivers’ willingness to implement new ACT strategies or attain their goals. The coach navigated these barriers through motivational conversations, supportive listening, and adaptable goal setting to accommodate individual preferences and limitations. Extrinsic barriers included time constraints, demanding caregiving responsibilities, and the evolving condition of the care recipient, leading to fluctuations in the caregivers’ schedules and diaries.

Regarding the nontherapeutic nature of the coaching, the coach expressed that “caregivers just needed to be heard.” However, providing motivation without knowing the context was often not possible, resulting in the duration of sessions exceeding the planned 20 minutes. Although the longer calls were still feasible for the coach, providing flexible availability was not always easy to adjust to daily work responsibilities. Furthermore, the coach noted that conducting data collection fully over the web and providing technical support posed occasional challenges that sometimes required contacting third parties and could result in delays.

**Discussion**

**Principal Findings**

This feasibility study was conducted in response to the demand for interventions that are both more scalable and personalized for family caregivers of people with dementia. The findings of this study revealed that caregivers acknowledged the beneficial impact of psychological flexibility. This recognition resulted in perceived positive effects not only in self-management and caregiving quality but also in the applicability of ACT in noncaregiving situations in daily life. This outcome is in line with that of previous research that ACT can be transdiagnostic in addressing psychological flexibility in a heterogeneous sample of caregivers who provided care for individuals at various stages and with various types of conditions, including dementia [46,47]. The findings of this study revealed a high adherence rate of 83% (24/29) at the postintervention measurements, which exceeded the average adherence rate of 57% found in self-guided ACT interventions [48], 73% found in self-help interventions [24], and 69% found in internet-based treatments for the general population [49]. High adherence and overall satisfaction are particularly important findings as family caregivers of people with dementia have the highest dropout rates and reported problems with access and usability of eHealth interventions [4].

The aim of this guided web-based intervention was to broaden the evidence base by exploring practical and conceptual barriers and facilitators influencing feasibility and acceptability. This investigation provides additional insights for refining future interventions and potentially facilitates the implementation of effective controlled trials on a larger scale. This web-based self-help intervention was designed for family caregivers of people with dementia and integrated ACT modules with collaborative goal setting based on personal values and weekly nontherapeutic coaching.

**Personalizing ACT Interventions Through Goal Setting**

In line with previous research, our findings showed that collaborative goal setting based on caregivers’ values holds promise as an approach to address the specific needs of caregivers and personalize interventions [36]. In addition, collaborative reflections and guidance played a crucial role in the process of value, need, and resource identification when
setting SMART goals as well as adjusting goals or timings throughout the intervention. This adaptability was perceived as necessary due to the continuously changing care demands and concurrent circumstances for caregivers. Multiple participants who set goals consistent with their originally selected values changed their goals focusing on a different value domain. Consistent with previous research, our findings underscore the significance of maintaining flexibility in future intervention designs, allowing for the selection of personally relevant values and value-focused exercises [50].

From the ACT perspective, values are meaningful qualities that cannot be obtained or finished but rather guide long-term patterns of behavior [11]. Subsequently, acknowledging personal values and nesting specific goals underneath them is more likely to drive effective behavior change in the long term [51]. Thus, the shift in focus to a different value domain may represent a potential intervention effect, enhancing caregivers’ motivation to align their lives with their values. Further investigation is also warranted to closely examine the factors that drive caregivers to modify their value-based goals following their involvement in a value-based intervention. Such an exploration will shed light on whether these changes in goals are motivated by a genuine desire of caregivers to align their lives with their core values.

In the context of personalizing ACT interventions, facilitating skill building and modifying behavioral responses in older adults with regard to potential concurrent challenging circumstances (eg, health conditions) might be especially important [52]. Moreover, technology offers opportunities to personalize future interventions by providing real-time feedback or guidance (eg, using experience sampling methodology and “just-in-time” interventions), promoting engagement with and managing multiple goals over time, simplifying complex goals, and facilitating collaborative or interactive reflections [53]. Personalizing eHealth interventions for future studies is particularly noteworthy as person-centered (family) care emerges as a crucial scope of research in the path of global dementia care [54].

Utility of ACT for Behavior Change in Caregivers

Our findings suggest that ACT contributes to an increase in perceived psychological flexibility and perspective taking. For some caregivers, increased psychological flexibility might enable them to prioritize focusing on personal growth and self-care as well as embracing challenges over acquiring high goal attainment scores or external validations and motivation (eg, those provided by the coach) [55]. However, perspective taking for some caregivers did not lead to a reported change in behavior.

For some caregivers, defining an appropriate goal or putting goals into action was challenging due to concurrent circumstances (eg, comorbidities). Family caregivers noted that, despite their change in attitude through practicing ACT, reconfiguring their habitual responses based on the new insights from ACT remained challenging in certain situations. It was difficult to adopt a new mindset and avoid reacting impulsively, especially in response to the unpredictable symptoms of the person with dementia. This suggests that behaviors may be triggered by contextual cues rather than being solely the result of mindful attention or personal willpower [56].

This finding is in line with that of previous research highlighting that habits might impact the relationship between attitudes and behavior, resulting in the regulation of desirable behavior change in the long term. Therefore, a change in attitude in caregivers, although crucial for behavior change, might be influenced by situational factors (eg, behavioral symptoms of the person with dementia), potentially making them less reliable predictors of behavior [56]. Addressing habit formation in the context of dementia caregiving is particularly noteworthy as behavioral symptoms and the deteriorative nature of dementia might lead to an unstable, impulsive, unpredictable, and stressful context for caregivers.

Future design of behavior change interventions can target habit formation and the interplay between attitudes and habits for the consolidation of effective upskilling and long-term behavior change in caregivers. This effort may involve developing ACT interventions with additional modules, such as behavior modification strategies for managing disruptive behaviors [13], tailored mindfulness, or habit reversal training [57]. In this endeavor, specific attention to discovering intervention mechanisms may benefit from the use of experience sampling methodology as a quantitative approach [58].

In addition, several participants in this study suggested the inclusion of an explicit module addressing the specific context of dementia and dementia caregiving. Accordingly, a module designed to incorporate ACT principles and provide guidance on managing dementia-specific challenges, such as effectively managing repetitive questions from people with dementia, could significantly enhance the applicability of ACT to their unique situation. This result is in line with those of previous research highlighting that caregiving itself is a natural value for caregivers [59]. The future development of ACT interventions for caregivers can be centered on values associated with caregiving (ie, maintaining care and improving the relationship with the person with dementia) while allowing for variations among different caregiver profiles based on their individual value preferences [28]. Further empirical research is now needed for further evidence-based understanding of the impact of value commitment on caregiver and care recipient outcomes.

Guided Self-Help Interventions for Caregivers

In the context of caregiving, adhering to web-based self-help interventions and incorporating them into everyday life commitments is thought to be improved by embedding more personal retention approaches (eg, telephone calls), provision of flexibility, personalization, scalability, and guidance [4,14,23]. Our study showed that a web-based self-help ACT intervention with synchronous motivational coaching was highly feasible and acceptable for family caregivers, suggesting that ACT and skill building can be learned through self-help [24].

It is important to recognize the coach’s significant role in enhancing adherence, motivation, and confidence in technology use as well as fostering a willingness to adopt eHealth among family caregivers of people with dementia. In line with previous research, our findings suggest that providing training on
technological features and access to troubleshooting might be beneficial for caregivers with lower digital literacy as, this way, they might be less likely to perceive eHealth as difficult, incompatible, or ineffective. Furthermore, including dedicated coaches for guidance and offering technical support might more likely generate a positive attitude toward eHealth [29].

Although this study provided insights into the role of coaching in maintaining intervention adherence, providing resource-heavy support does not reflect how ACT-based programs are often used at a broader scale and outside the research context [48]. Future designs of personalized and guided eHealth interventions can explore how technology can provide additional support to health care professionals as providers to reduce personal resources and also facilitate the design of effective and implementable interventions on a large scale. Subsequently, guided self-help interventions for family caregivers can evaluate the effects of varying degrees of contact, allowing for the realization of what is the minimum amount of contact that should accompany self-help to obtain the maximum benefit [60]. To support effective, acceptable, and sustainable interventions, future approaches could involve caregivers in the process of designing interventions to capture both the context and dementia-specific needs of caregivers to be tested in larger samples [9].

**Adapting ACT to a Web-Based Format for Older Adults**

Although websites are the most common way of delivering web-based ACT interventions and are typically more feasible for researchers to build [61], our results showed the preference of some participants to use the intervention via their mobile phones (web application). Considering that mobile apps are more accessible and are the most commonly accessed method of self-help, future research can consider ACT-based apps for family caregivers to address the discrepancy between the ACT literature and real-world practice [48]. This technological development is noteworthy for bridging the gap to make therapy available, accessible, and affordable for larger populations of family caregivers who do not need heavy support [62]. Implementing interventions in the real world is essential for caregivers as, despite the need for behavior change interventions, most feasible and effective eHealth interventions for family caregivers of people with dementia are not yet ready for implementation and, thus, not implemented in the real world [63].

**Strengths and Limitations**

This study introduced a blended intervention to address the crucial need for effective interventions in dementia caregiving by exploring the feasibility of a web-based ACT intervention tailored for this population. The strengths of this study lie in its contribution to the need for further evidence-based interventions in the area of ACT and family care, paving the way for future controlled trials and intervention refinement. Additional strengths include high adherence rates, flexibility in delivering the intervention to a diverse population of caregivers of people with dementia, and shedding light on barriers and facilitators that family caregivers experienced over the course of the guided web-based self-help intervention. The mixed methods approach combined quantitative adherence rates and qualitative caregiver feedback for a comprehensive understanding of the intervention.

Nevertheless, this study might be influenced by potential biases, and when interpreting the findings, it is crucial to acknowledge its limitations. Although conducting feasibility studies before an RCT can ensure the design of studies with a higher likelihood of success, the small sample size in this study necessitates caution when generalizing the results to a larger population of caregivers of people with dementia. The study sample size was relatively small (N=30), which may limit the generalizability of the findings. Larger samples are required to draw more robust conclusions and account for potential individual differences among caregivers. In addition, the study duration was 9 weeks, and it lacked a control group, which made it challenging to determine whether the observed positive outcomes were solely due to the intervention or whether other factors may have contributed to the results. Conducting larger studies with long follow-up assessments and including a control group would provide a better basis for assessing this guided web-based intervention.

Participation in the study was voluntary and occurred over the web. Thus, individuals who chose to participate may not constitute a fully representative sample, potentially skewing toward those who are more technologically savvy with higher levels of education and greater familiarity with technology. It is crucial to acknowledge this selection bias and consider generalizability concerns when interpreting the findings. ACT principles underscore that individuals do not always have the autonomy to select the content of any given situation. To convey this concept, ACT uses metaphors that may necessitate intellectual engagement and abstract reasoning [10]. Considering the above-average level of education among the study participants and the availability of a trained coach for questions, our findings may not be broadly generalizable to caregivers with lower digital literacy or a lower educational background.

We used a guided ACT-based intervention blended with other non-ACT complementary techniques (eg, goal setting) and components (eg, weekly coaching), which might have had therapeutic effects. Specifically, some caregivers received more than the anticipated 20 minutes of coaching, leading to an unequal distribution of coaching among participants. This variability, as well as the use of self-report questionnaires, could potentially influence the generalizability of the study findings. In addition, most caregivers were at an early stage of caregiving and cared mainly for people with Alzheimer disease, limiting the understanding of how the intervention works across different contexts and stages of dementia caregiving.

**Conclusions**

This study was conducted to evaluate the feasibility and acceptability of a guided web-based self-help ACT intervention for family caregivers of people with dementia. The high adherence rate and positive feedback from caregivers indicate the intervention’s feasibility and acceptability. The findings suggest that family caregivers can learn ACT principles and use them to enhance their psychological flexibility through self-help. Moreover, personalizing the intervention through collaborative goal setting based on individuals’ values was found to be
promising for addressing the specific needs of caregivers. The findings also suggest that ACT can be adapted to a web-based format, increasing accessibility and scalability for a diverse and large sample of caregivers. However, the absence of a control group and small sample size limit the drawing of definitive conclusions. Some caregivers faced challenges in translating new skills into behavior change due to habitual responses. Larger controlled trials are needed to validate the feasibility in a more diverse caregiver sample and determine the effectiveness of ACT-based interventions in this population. This study highlights the need for future interventions to address habit formation and the interplay between attitudes and habits in unpredictable and continuously changing caregiving contexts. In addition, exploration of the impact of value-based behavior on caregiver and care recipient outcomes should be considered in further research. It is hoped that the results of this feasibility study will pave the way for future effective controlled trials and the implementation of evidence-based research in real-world settings.

Acknowledgments
The authors express their sincere gratitude to all family caregivers who participated in this research, acknowledging their enthusiasm and willingness to share their experiences. We also appreciate the presence of Roos Roberts for their dedication during screening and coaching sessions, which has made valuable contributions to this study. The research presented in this paper received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement 813196 and was carried out within the framework of the Dementia: Intersectorial Strategy for Training and Innovation Network for Current Technology (DISTINCT) Innovative Training Network.

Conflicts of Interest
None declared.

Multimedia Appendix 1
CONSORT-EHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online Telehealth) checklist V 1.6.1.
[PDF File (Adobe PDF File), 1088 KB - aging_v7i1e53489_app1.pdf ]

Multimedia Appendix 2
An overview of family caregivers’ potential personal values.
[DOCX File, 15 KB - aging_v7i1e53489_app2.docx ]

Multimedia Appendix 3
Steps of collaborative goal setting.
[DOCX File, 531 KB - aging_v7i1e53489_app3.docx ]

Multimedia Appendix 4
The interface of the website used to deliver the intervention.
[DOCX File, 2291 KB - aging_v7i1e53489_app4.docx ]

Multimedia Appendix 5
An overview of the weekly coaching questions.
[DOCX File, 17 KB - aging_v7i1e53489_app5.docx ]

Multimedia Appendix 6
An overview of the Program Participation Questionnaire.
[DOCX File, 25 KB - aging_v7i1e53489_app6.docx ]

Multimedia Appendix 7
Semistructured interview with the coach.
[DOCX File, 20 KB - aging_v7i1e53489_app7.docx ]

Multimedia Appendix 8
An overview of the acceptability of the self-help modules.
[DOCX File, 19 KB - aging_v7i1e53489_app8.docx ]

https://aging.jmir.org/2024/1/e53489
References


61. Stawarz K, Freist C, Coyle D. Use of smartphone apps, social media, and web-based resources to support mental health and well-being: online survey. JMIR Ment Health 2019 Jul 12;6(7):e12546 [FREE Full text] [doi: 10.2196/12546] [Medline: 31301126]


Abstract

Background: The increasing prevalence of Alzheimer disease and Alzheimer disease–related dementia in the United States has amplified the health care burden and caregiving challenges, especially for caregivers of people living with dementia. A web-based care planning tool, Olera.care, was developed to aid caregivers in managing common challenges associated with dementia care.

Objective: This study aims to preliminarily evaluate the quality and usability of the Olera.care platform and assess the preferences of using the technology and interests in learning about different older adult care services among caregivers.

Methods: For interview 1, we aim to understand caregiving needs and let the participants start engaging with the platform. After they engage with the platform, we schedule the second interview and let the participants complete the Mobile Application Rating Scale. The survey also included sociodemographic characteristics, caregiving experiences, communication preferences in technology adoption, and older adult care service use and interests. Descriptive statistics were used to describe the quality and usability of the platform and characteristics of the participants. We conducted 2-sample 2-tailed t tests to examine the differences in the Mobile Application Rating Scale evaluation scores by caregiver characteristics.

Results: Overall, 30 adult caregivers in Texas completed the evaluation. The majority were aged ≥50 years (25/30, 83%), women (23/30, 77%), White (25/30, 83%), and financially stable (20/30, 67%). The Olera.care platform evaluation showed high satisfaction, with an overall mean rating of 4.57 (SD 0.57) of 5, and scored well in engagement (mean 4.10, SD 0.61), functionality (mean 4.46, SD 0.44), aesthetics (mean 4.58, SD 0.53), and information quality (mean 4.76, SD 0.44) consistently across all participants. A statistically significant difference (P=.02) was observed in functionality evaluation scores by duration of caregiving, with caregivers dedicating more hours to care rating it higher than those providing less care (mean 4.6, SD 0.4 vs mean 4.2, SD 0.5). In addition, caregivers with less caregiving experience reported significantly higher evaluation scores for aesthetics (P=.04) and information quality (P=.03) compared to those with longer years of caregiving. All participants expressed a willingness to recommend the app to others, and 90% (27/30) rated the app overall positively. Most of the participants (21/30, 70%) favored anonymous interactions before receiving personalized feedback and preferred computer browsers over mobile apps. Medical home health services were the most used, with a diverse range of services being used. Caregiver support groups, medical providers, memory care, meal services, and adult day care were among the most desired services for future exploration.
**Conclusions:** The Olera.care web-based platform is a practical, engaging, easy-to-use, visually appealing, and informative tool for dementia caregivers. Future development and research are essential to enhance the platform and comprehensively evaluate it among a broader population.

(JMIR Aging 2024;7:e55132) doi:10.2196/55132

**KEYWORDS**

evaluation; usability; family caregiver; Alzheimer disease; dementia; digital health; mobile phone

**Introduction**

**Background**

With the population of Americans living with Alzheimer disease (AD) and AD-related dementia projected to grow from 6.7 million to 13.8 million by 2060, there is a need for effective and innovative solutions to address the increasing health care burden, particularly with caregivers of people living with dementia [1,2]. Oftentimes, these caregivers are unpaid or informal family members or friends who provide care mostly related to activities of daily living [3,4]. Due to the variability in disease progression and the caregiver’s personal needs, family caregivers encounter diverse unmet needs, including challenges related to the physical and emotional deterioration of the people living with dementia [5-7]. Particularly, caregivers of people living with dementia report difficulties related to assisting the care recipient with activities of daily living, identifying the right older adult living services, navigating the financial and legal aspects of caregiving for people living with dementia, and finding relevant and concise information on dementia and dementia caregiving [5-7]. Due to the multifaceted and individualized burdens that caregivers of people living with dementia face, they are particularly susceptible to experiencing emotional, physical, and financial challenges that increase with disease progression [3,6,7].

To aid informal caregivers with challenges related to dementia care, many digital technologies have been developed that focus on either the needs of the caregiver or of the people living with dementia [8]. These solutions range from web-based training to web-based forums and caregiving groups, psychological and educational forums, and videoconferencing technologies [9,10]. Despite the many digital interventions to aid caregivers and their care recipients, these solutions do not adequately address the individual needs of caregivers of people living with dementia, leading to limited adoption of technologies outside of pilot studies [5,11]. Previous studies have shown that, despite the availability of digital interventions, the unique and individual needs of caregivers of people living with dementia are not always adequately addressed, leading to limited technology adoption and potentially contributing to sustained or increased emotional, physical, and financial burdens associated with caregiving [6,9]. To increase the use of technology that aids caregivers of people living with dementia, technology development should involve caregivers’ feedback on usability and align with caregivers’ expectations and needs [7,12,13]. However, not many studies have investigated the usability and usefulness of technology interventions for dementia care or involved caregivers of people living with dementia in the development process to accurately address caregivers’ needs [10,12,14-16].

With the plethora of digital technologies available today to aid in dementia caregiving, there is a need for an evidence-based, engaging, adaptable, and preference-based platform for informal caregivers of people living with dementia to aid them in identifying resources and education on dementia relevant to their and their care recipient’s needs [5,7,11,17]. Previously, our qualitative study reported that caregivers are looking for several features in such a web-based navigation and resource platform: “(1) a comprehensive database of commonly needed professional services, (2) mental health and caregiver support groups, (3) educational resources on dementia and caregiving, [and] (4) a platform that is easy to use, aesthetic, reliable, and interactive” [5]. Some current solutions that function as resource finders or navigators include the Community Resource Finder by AARP and the Alzheimer’s Association, the Alzheimer’s Navigator by the Alzheimer’s Association, and CareNav by the Family Caregiver Alliance. While these web-based tools provide databases and tips for locating resources and education on dementia and caregiving, they do not provide categorized recommendations or resources tailored to the stage of dementia of the people living with dementia and based on caregiver characteristics and preferences.

The overall performance rating of a digital app by users is greatly affected by several behavioral factors according to the theory of planned behavior, which states that an individual’s intention to use or adopt technologies is greatly influenced by their attitude toward the technologies [18,19]. Thus, the more favorably a caregiver views a technological intervention, the more likely they will be to adopt and use the product. In addition, an individual’s adoption of a technological intervention is greatly influenced by their perception of technology self-efficacy, which is defined as an individual’s confidence in applying a technology to perform a task [20-23]. The more positive an individual’s attitude and initial experience, the more likely they are to use the technology as their perceived self-efficacy increases [24]. The Mobile Application Rating Scale (MARS) is a widely used tool to test the performance of health mobile apps and digital platforms based on the functionality, design, information quality, engagement, and subjective quality of the digital apps [25].

**Objectives**

To address caregivers’ needs and expectations of a web-based platform to assist them in the care of their care recipient, we developed a web-based care planning tool, Olera.care, that assists caregivers in navigating common challenges by supplying personalized recommendations and curated sets of resources.
(eg, care services, products, and professionals) as well as education on dementia and caregiving for people living with dementia. This pilot study aims to evaluate the functionality and usability of the initial Olera.care platform for caregivers of people living with dementia in addressing their needs with personalized education and resource matching.

**Methods**

**Overview**

This pilot study was conducted to support the development and evaluation of the usability of the Olera.care digital platform. The development and testing of the platform adopt an iterative “build-measure-learn” approach (Figure 1) that places caregivers at the forefront of design and development, ensuring that our platform continually evolves to meet their evolving needs and expectations [26]. This framework ensures that our platform iteratively evolves through an ongoing process of design and development, incorporating frequent touch points with family caregivers to assess the usability and functionality of a given prototype and align subsequent development with caregiver wishes and expectations for digital assistance technology. Participants in the study were engaged in 2 rounds of Zoom (Zoom Video Communications, Inc) or telephone interviews from January 2022 to May 2022 to (1) understand their caregiver needs, (2) engage with the platform, and (3) complete a technology survey assessment, including the modified MARS, via a Qualtrics (Qualtrics International Inc) web-based form. Each MARS item used a 5-point scale to assess the engagement, functionality, aesthetics, and information quality of the Olera.care digital platform among unpaid caregivers of people living with dementia in Texas.

Figure 1. The build-measure-learn approach used to research and develop Olera.care.

**Platform Development and Testing**

**Development of a Digital Platform Capable of Providing Personalized Information on the Legal, Financial, and Estate Planning Aspects of Caregiving for Dementia, Including Information on Local Resources Most Relevant to Caregivers**

We developed a robust web application that provides tailored educational videos and articles on topics associated with the most prominent challenges and struggles that caregivers of people living with dementia face as evidenced by our previous work identifying common pain points in the caregiving journey [5]. Personalized information is curated based on answers to a caregiving questionnaire and an algorithm developed to sift through a data repository and present the most relevant information pertaining to a user’s circumstances. Our growing content repository currently hosts 66 original articles and video postings that cover various topics, including legal, financial, and estate planning. In addition to personalized education, the developed web application can present tailored listings of relevant professionals in the legal, financial, home care, older adult living, and older adult care coordination industry (Figure 2). Credentialed professionals are presented in a personalized directory that is curated based on our algorithm’s assessment of caregivers’ current professional needs, preferences, and geographic location.
Compilation and Management of Relevant Financial, Legal, and Estate Planning Resources Available in Our Target Area

We have created a repository database that includes information on >22,000 professional service providers and 66 original educational articles or videos. This database is integrated with our web application’s user interface to allow users to sort through it quickly and effortlessly for information most relevant to their current needs (Figure 3). Of note, relevant service providers included in our database are certified financial planners, Medicare insurance agents, older adult law attorneys, older adult care planners, older adult housing facilities (memory care and assisted or independent living), rehabilitation centers, hospice or palliative care agencies, and in-home professional caregivers as well as home health services. These services have been included because of direct commentary noted on interviewing current dementia caregivers about their service needs when providing AD and AD-related dementia family caregiving for a loved one [5].

Figure 2. Content repository example of the Olera.care platform.

Figure 3. Personalized dashboard for older adult care planning on the Olera.care platform.
Preliminary Assessments of the Usability and Functionality of the Technology Using Techniques Such as End-User Surveys and Real-Time Monitoring

We evaluated the usability and functionality of our caregiver support platform with a pilot panel of 30 current family caregivers of persons living with dementia who met our eligibility criteria. Participants interacted with the web application in a test environment. They were asked to use the main functions and rate the usability and functionality on a modified MARS, which is a well-established instrument used as a benchmark for digital health apps. In addition, real-time monitoring during participation helped researchers identify areas for improvement in the user experience.

Participant Eligibility and Recruitment

Participants were recruited through academic and community networks, such as the Texas A&M University Center for Community Health and Aging and the Brazos Valley Area Agency on Aging. We used both traditional and digital channels to reach out to potential participants. These included printed materials such as physical or electronic flyers, outreach emails, and social media platforms (eg, Facebook, Instagram, and Twitter) as well as engagement with web-based forums such as local caregiver support group meetings.

Individuals expressing interest in participating in the study were provided 2 options to complete the eligibility screening: participants could access a web-based form by following the recruitment materials’ QR code or web link; alternatively, they could contact the Olera.care team via telephone to obtain information about the study and take the eligibility screening survey over the telephone.

The eligibility screening survey collected relevant information to examine whether interested individuals met the following inclusion criteria: (1) be aged ≥18 years; (2) be a nonpaid caregiver of a person living with dementia; (3) be the adult child, spouse or partner, other family member, or legal guardian of a person living with dementia; (4) be engaged in making legal or financial, older adult living, or medical decisions for a person living with dementia; (5) be seeking older adult care services in Texas; and (6) have access to a smartphone or computer with internet access.

After the eligibility screening process, participants who met the inclusion criteria were asked to provide their full name, contact information, preferred mode of contact (ie, telephone or email), and preferred day and time for subsequent outreach by the research team.

Ethical Considerations

Ethics approval was obtained from the Texas A&M University Institutional Review Board (2021-0943D). The study personnel asked all participants to provide electronic informed consent in the screening survey. Upon the completion of the eligibility screening survey, participants were presented with an informed consent document. This document provided instructions on how to convey their consent to participate in the study and their willingness to be recorded and followed up via the Qualtrics platform. The consent form covered important information, including the rationale for inclusion; the research objectives; the voluntariness of participation with the option to withdraw at any point; the anticipated participation duration and procedures; the potential risks, benefits, and costs of participation; and how participant confidentiality would be protected.

Assessments and Measurements

Evaluating the Quality of the Olera.care Platform Using the MARS

We used the MARS as a robust assessment tool to evaluate the quality of the Olera.care platform among the caregiver participants. The MARS provides a multidimensional assessment of the engagement, functionality, aesthetics, information quality, and overall subjective quality of the Olera.care platform. To tailor the assessment to our study’s specific context, we selected 13 items from the MARS questionnaire that were directly relevant to the Olera.care platform. Our methodology for adapting the MARS questionnaire was 2-pronged: first, we evaluated and retained items based on their relevance to our platform’s functions, omitting nonapplicable elements such as gestural design; and second, we modified the wording of the retained items to better reflect our platform’s unique features. These items were adapted and modified for our assessment objectives while retaining the original item classification across the 5 dimensions. The modified items and responses, while maintaining the integrity of the MARS, ensure an effective evaluation of our digital platform. Participants provided ratings for each MARS item using a 5-point scale (1=ineffective, 2=poor, 3=acceptable, 4=good, and 5=excellent), with each response tailored to the content of the respective item. An overall subjective quality rating of ≥3.6 was set as the threshold to indicate good usability and quality of the Olera.care platform according to past literature [27], allowing us to effectively report the platform’s overall quality from assessment among the caregiver users.

Assessing Willingness, Self-Efficacy, and Communication Preferences in Technology Adoption

To understand the willingness and preference for technology use, we further assessed participants’ intention to use the Olera.care platform, self-efficacy for using the technology, and preferences for web-based communication and platform format. The participants were asked whether they would like to use the technology in the future, with the response options being “yes,” “maybe,” and “no.” They were also asked about their confidence level in using the technology, with the response options being “uncertain,” “neither certain nor uncertain,” “somewhat confident,” and “very confident.” In addition, participants were asked about their preference for anonymity when seeking information on the internet and whether they would consider sharing personal information for receiving individualized answers. The net promoter score, which was created in 2003 and has been used in a variety of industries such as insurance, technology services, communications, and health care [28], was used to evaluate the willingness of participants and provide insights for user experience management. Responses on a scale ranging from 0=strongly disagree to 10=strongly agree were used. Scores from 0 to 6 were classified as detractor, scores of
7 and 8 were categorized as \textit{passive}, and scores of 9 and 10 were designated as \textit{promoter}. Participants were also asked whether they preferred a website-based format or a mobile app format.

\textbf{Assessing the Use of, and Interests in, Older Adults Care Services Among Caregivers}

To gain a deeper understanding of our targeted population’s preferences and needs for planning older adult care and to continually enhance and optimize our platform, we assessed the level of interests in older adult care services among caregivers. Caregivers were presented with 21 types of older adult care services (e.g., home health, hospice care, memory care, caregiver support group, and assisted living), and they were asked to select \(\geq 1\) of the following responses: “currently using,” “have used before,” “would like to learn more,” and “would never use.” This assessment allowed us to gather valuable insights into caregivers’ engagement and interests in various older adult care options, informing our efforts to better serve their needs.

\textbf{Sociodemographic and Caregiving Characteristics of Caregivers}

We collected the sociodemographic and caregiving characteristics of participants to understand how representative our study population was for the Texas caregiver profile and whether the major platform evaluation outcomes differ by the background characteristics of participants. The caregiver characteristics collected included age, sex, race, ethnicity, the highest level of education completed, employment status, general financial status, caregiving role, and length of providing care.

\textbf{Analysis}

Descriptive statistics were used to describe the engagement, functionality, aesthetics, information quality, and subjective quality of the platform and caregiver characteristics and responses. Mean scores and SDs were calculated for each MARS item. We conducted 2-sample 2-tailed \(t\) tests to compare the differences in the major MARS evaluation scores by caregiver characteristics. All analyses were conducted using Stata (version 17.0; StataCorp LLC).

\textbf{Results}

\textbf{Participants’ Characteristics}

Of the initial 822 respondents who completed the prescreening surveys, 150 (18.2\%) met the eligibility criteria, of whom, after excluding 115 (76.7\%) individuals for not being available to attend interviews, 35 (23.3\%) were enrolled into the study. Of these 35 enrolled individuals, 30 (86\%) interacted with the platform and completed the study survey of technology evaluation and caregiving needs (Figure 4). The sociodemographic and caregiving characteristics of participants are summarized in Table 1.

Of the 30 participants, the majority were aged \(\geq 50\) years (n=25, 83\%), women (n=23, 77\%), White (n=25, 83\%), non-Hispanic (n=22, 73\%), and were employed for wages (n=12, 40\%) or retired (n=12, 40\%). Financially, 67\% (20/30) had surplus funds at the end of each month, while 30\% (9/30) just about managed to meet their expenses or faced deficits. Most of the participants were recruited via email invitations (13/30, 43\%) or web-based advertisements (11/30, 37\%), with some also recruited through in-person presentations and personal connections (5/30, 17\%). In terms of caregiving characteristics, the majority of the participants identified themselves as primary caregivers (20/30, 67\%), reported to have provided care for at least 1 year (28/30, 93\%), and dedicated at least 20 hours weekly to caregiving in the past 3 months (21/30, 70\%).
Figure 4. Study flow diagram.

Enrollment

Assessed for eligibility (n=822)

Excluded because not meeting inclusion criteria (n=672)

Eligible (n=150)

Excluded because not available to attend interviews (n=115)

Completed Interviews (n=35)

Excluded because did not complete survey (n=5)

Completed survey (n=30)

Follow-up

Analysis

Analyzed (n=30)
Table 1. Sociodemographic and caregiving characteristics of participants (n=30).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Participants, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td></td>
</tr>
<tr>
<td>35-49</td>
<td>5 (17)</td>
</tr>
<tr>
<td>50-64</td>
<td>12 (40)</td>
</tr>
<tr>
<td>≥65</td>
<td>13 (43)</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Female</td>
<td>23 (77)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Asian, Native Hawaiian, or Pacific Islander</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>2 (7)</td>
</tr>
<tr>
<td>White</td>
<td>25 (83)</td>
</tr>
<tr>
<td>Multiracial</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Spanish, Hispanic, or Latinx origin or descent</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Other</td>
<td>27 (90)</td>
</tr>
<tr>
<td>Highest level of education completed</td>
<td></td>
</tr>
<tr>
<td>Some college but no degree</td>
<td>5 (17)</td>
</tr>
<tr>
<td>Associate degree</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Bachelor’s degree</td>
<td>8 (27)</td>
</tr>
<tr>
<td>Graduate degree</td>
<td>14 (47)</td>
</tr>
<tr>
<td>Employment status</td>
<td></td>
</tr>
<tr>
<td>Employed for wages</td>
<td>12 (40)</td>
</tr>
<tr>
<td>Homemaker or self-employed</td>
<td>4 (13)</td>
</tr>
<tr>
<td>Unemployed or unable to work</td>
<td>2 (7)</td>
</tr>
<tr>
<td>Retired</td>
<td>12 (40)</td>
</tr>
<tr>
<td>General financial status at the end of the month</td>
<td></td>
</tr>
<tr>
<td>End up with some money left over</td>
<td>20 (67)</td>
</tr>
<tr>
<td>Have just enough to make ends meet</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Not have enough money to make ends meet</td>
<td>2 (7)</td>
</tr>
<tr>
<td>Do not know</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Caregiving role in providing care for ≥1 adults aged &gt;50 y</td>
<td></td>
</tr>
<tr>
<td>Participant is the primary caregiver</td>
<td>20 (67)</td>
</tr>
<tr>
<td>Someone else is the primary caregiver</td>
<td>4 (14)</td>
</tr>
<tr>
<td>The participant shared caregiving responsibilities about equally with someone else</td>
<td>5 (17)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (3)</td>
</tr>
<tr>
<td>How long has the participant been providing care or assistance for the care recipients? (y)</td>
<td></td>
</tr>
<tr>
<td>0.5 to &lt;1</td>
<td>2 (7)</td>
</tr>
<tr>
<td>1 to &lt;5</td>
<td>14 (47)</td>
</tr>
<tr>
<td>5 to &lt;10</td>
<td>8 (27)</td>
</tr>
<tr>
<td>≥10</td>
<td>6 (20)</td>
</tr>
<tr>
<td>Over the past 3 months, about how many h/wk has the participant provided some form of care for ≥1 adults aged &gt;50 y?</td>
<td></td>
</tr>
</tbody>
</table>
Quality Assessment of the Olera.care Platform

Overview
The descriptive statistics of the Olera.care platform evaluation by the MARS dimensions and items are presented in Table 2. The mean overall satisfaction rating of the Olera.care platform among the participants was 4.57 (SD 0.57) of 5. Mean scores for engagement, functionality, aesthetics, and information quality were 4.10 (SD 0.61), 4.46 (SD 0.44), 4.58 (SD 0.53), and 4.76 (SD 0.44), respectively.
Table 2. Modified Mobile Application Rating Scale items and responses of participants.

<table>
<thead>
<tr>
<th>Dimensions, items, questions, and response items</th>
<th>Participants, n (%)</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Engagement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Entertainment: compared to other older adult care finders, websites, or alternative resources for older adult care information you may have used, do you think this app is fun to use?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highly entertaining</td>
<td>5 (17)</td>
<td>3.67 (0.80)</td>
</tr>
<tr>
<td>Fun</td>
<td>11 (37)</td>
<td></td>
</tr>
<tr>
<td>OK, fun enough</td>
<td>13 (43)</td>
<td></td>
</tr>
<tr>
<td>Mostly dull</td>
<td>1 (3)</td>
<td></td>
</tr>
<tr>
<td>Dull</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td><strong>Interestingness</strong>: compared to other older adult care finders, websites, or alternative resources for older adult care information you may have used, does this app present its content in an interesting way?</td>
<td></td>
<td>4.53 (0.57)</td>
</tr>
<tr>
<td>Very interesting</td>
<td>17 (57)</td>
<td></td>
</tr>
<tr>
<td>Interesting</td>
<td>12 (40)</td>
<td></td>
</tr>
<tr>
<td>Slightly interesting</td>
<td>1 (3)</td>
<td></td>
</tr>
<tr>
<td>Mostly uninteresting</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Not interesting</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td><strong>Functionality</strong></td>
<td></td>
<td>4.46 (0.44)</td>
</tr>
<tr>
<td>Performance: how well do the app’s features, components, and buttons work?</td>
<td></td>
<td>4.57 (0.57)</td>
</tr>
<tr>
<td>Works perfectly</td>
<td>18 (60)</td>
<td></td>
</tr>
<tr>
<td>Very functional</td>
<td>11 (37)</td>
<td></td>
</tr>
<tr>
<td>App works overall</td>
<td>1 (3)</td>
<td></td>
</tr>
<tr>
<td>Some functions work</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>App is broken</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td><strong>Ease of use</strong>: how easy is it to learn how to use the app?</td>
<td></td>
<td>4.57 (0.50)</td>
</tr>
<tr>
<td>Very simple</td>
<td>17 (57)</td>
<td></td>
</tr>
<tr>
<td>Easy to learn</td>
<td>13 (43)</td>
<td></td>
</tr>
<tr>
<td>Usable</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Somewhat confusing</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Very confusing</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td><strong>Navigation</strong>: does moving between pages make sense?</td>
<td></td>
<td>4.28 (0.59)</td>
</tr>
<tr>
<td>Moving between pages is perfectly logical, easy, clear and intuitive</td>
<td>10 (33)</td>
<td></td>
</tr>
<tr>
<td>Moving between pages is easy to understand and navigate</td>
<td>17 (57)</td>
<td></td>
</tr>
<tr>
<td>Moving between pages is understandable after some time and effort</td>
<td>2 (7)</td>
<td></td>
</tr>
<tr>
<td>Moving between pages is understandable after a lot of time and effort</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Moving between pages is difficult</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>1 (3)</td>
<td></td>
</tr>
<tr>
<td><strong>Aesthetics</strong></td>
<td></td>
<td>4.58 (0.53)</td>
</tr>
<tr>
<td>Visual appeal: how good does the app look?</td>
<td></td>
<td>4.53 (0.63)</td>
</tr>
<tr>
<td>Very visually appealing</td>
<td>18 (60)</td>
<td></td>
</tr>
<tr>
<td>High-level visual appeal</td>
<td>10 (33)</td>
<td></td>
</tr>
<tr>
<td>Some visual appeal</td>
<td>2 (7)</td>
<td></td>
</tr>
<tr>
<td>Little visual appeal</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>No visual appeal</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td><strong>Graphics</strong>: how high is the quality of graphics, buttons, and content?</td>
<td></td>
<td>4.53 (0.63)</td>
</tr>
</tbody>
</table>
### Dimensions, items, questions, and response items

<table>
<thead>
<tr>
<th>Description</th>
<th>Participants, n (%)</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very high quality</td>
<td>18 (60)</td>
<td></td>
</tr>
<tr>
<td>High quality</td>
<td>10 (33)</td>
<td></td>
</tr>
<tr>
<td>Moderate quality</td>
<td>2 (7)</td>
<td></td>
</tr>
<tr>
<td>Low quality</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Very poor quality</td>
<td>0 (0)</td>
<td></td>
</tr>
</tbody>
</table>

#### Layout: how would you rate the design? Are the arrangement and size of buttons and content on the screen appropriate?

<table>
<thead>
<tr>
<th>Description</th>
<th>Participants, n (%)</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very professional</td>
<td>22 (73)</td>
<td>4.67 (0.61)</td>
</tr>
<tr>
<td>Mostly professional</td>
<td>6 (20)</td>
<td></td>
</tr>
<tr>
<td>Satisfactory</td>
<td>2 (7)</td>
<td></td>
</tr>
<tr>
<td>Bad design</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Very bad design</td>
<td>0 (0)</td>
<td></td>
</tr>
</tbody>
</table>

#### Information

#### Accuracy of app description: after reviewing the home page, does the app contain what is advertised and described?

<table>
<thead>
<tr>
<th>Description</th>
<th>Participants, n (%)</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highly accurate</td>
<td>23 (77)</td>
<td>4.76 (0.44)</td>
</tr>
<tr>
<td>Mostly accurate</td>
<td>6 (20)</td>
<td></td>
</tr>
<tr>
<td>Somewhat accurate</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Slightly misleading</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Very misleading</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>1 (3)</td>
<td></td>
</tr>
</tbody>
</table>

#### Quality of information: is the content in the app relevant to helping with older adult care planning?

<table>
<thead>
<tr>
<th>Description</th>
<th>Participants, n (%)</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Highly relevant</td>
<td>22 (73)</td>
<td>4.76 (0.44)</td>
</tr>
<tr>
<td>Relevant</td>
<td>7 (23)</td>
<td></td>
</tr>
<tr>
<td>Moderately relevant</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Barley relevant</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Irrelevant content</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Missing</td>
<td>1 (3)</td>
<td></td>
</tr>
</tbody>
</table>

#### Quality of visual information: are images, videos, and graphics clear and easily understandable?

<table>
<thead>
<tr>
<th>Description</th>
<th>Participants, n (%)</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perfectly clear</td>
<td>22 (73)</td>
<td>4.73 (0.45)</td>
</tr>
<tr>
<td>Mostly clear</td>
<td>8 (27)</td>
<td></td>
</tr>
<tr>
<td>Somewhat clear</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Mostly unclear</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>Completely unclear</td>
<td>0 (0)</td>
<td></td>
</tr>
</tbody>
</table>

#### Subjective quality

#### Stimulates repeat use: how many times do you think you would use this app in the next 12 months if it was relevant to you?

<table>
<thead>
<tr>
<th>Description</th>
<th>Participants, n (%)</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;50</td>
<td>6 (20)</td>
<td>3.77 (0.86)</td>
</tr>
<tr>
<td>10-50</td>
<td>13 (43)</td>
<td></td>
</tr>
<tr>
<td>3-10</td>
<td>9 (30)</td>
<td></td>
</tr>
<tr>
<td>1-2</td>
<td>2 (7)</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>0 (0)</td>
<td></td>
</tr>
</tbody>
</table>

#### Worth recommending: would you recommend this app to people who might benefit from it?

<table>
<thead>
<tr>
<th>Description</th>
<th>Participants, n (%)</th>
<th>Score, mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I would recommend this app to everyone</td>
<td>13 (43)</td>
<td>4.27 (0.14)</td>
</tr>
</tbody>
</table>
Dimensions, items, questions, and response items | Participants, n (%) | Score, mean (SD)
--- | --- | ---
There are many people I would recommend this app to | 12 (40) | 4.57 (0.57)
There are several people I would recommend it to | 5 (17) |
There are a few people I would recommend this app to | 0 (0) |
I would not recommend this app to anyone | 0 (0) |
Overall satisfaction rating: what is your overall star rating of the app? (1 star=poor; 5 stars=excellent) | 4.57 (0.57)
5 | 17 (57) |
4 | 10 (33) |
3 | 1 (3) |
2 | 0 (0) |
1 | 0 (0) |
Missing | 2 (7) |

**Engagement**

Participants reported high levels of engagement with the Olera.care platform, with 97% (29/30) describing it highly entertaining or fun or fun enough, and 97% (29/30) expressing that it was interesting to interact with.

**Functionality**

In terms of functionality, the majority of the participants assessed the Olera.care platform positively, with 60% (18/30) perceiving it as working perfectly and 37% (11/30) rating it as very functional. Moreover, 57% (17/30) found it very simple to learn to use, and 43% (13/30) considered it easy to learn. An impressive 90% (27/30) of the participants reported that it was easy to navigate between pages.

**Aesthetics**

The aesthetics of the platform received favorable feedback from participants, with 93% (28/30) expressing that it was visually appealing, highlighting the high quality of graphics, buttons, and content. The design and layout of the content were described as professional.

**Information**

In terms of information quality, nearly all participants (29/30, 97%) noted that the app contained relevant and clear information. A substantial 73% (22/30) found the information highly relevant for older adult care planning, while 23% (7/30) considered it relevant. Visual information was deemed perfectly clear by 73% (22/30) of the participants and mostly clear by 27% (8/30).

**Subjective Quality**

Participants expressed a strong inclination to use the app in in the next 12 months, with 20% (6/30) planning to use it >50 times and 43% (13/30) aiming to use it between 10 and 50 times. Importantly, all participants indicated their willingness to recommend the app to those who may benefit from it. In terms of overall satisfaction, most of the participants rated the Olera.care platform positively, with 57% (17/30) rating it as excellent and 33% (10/30) as very good.

**Intention to Use the Olera.care Platform, Self-Efficacy of Technology Use, and Platform Preference**

In this study, a substantial proportion of the participants expressed a positive intention to use the Olera.care platform (Table 3), with 90% (27/30) indicating a definite willingness (“yes”), while the remaining 10% (3/30) expressed a more tentative interest (“maybe”). In terms of self-efficacy in using technology, 40% (12/30) of the participants reported feeling “very confident,” 57% (17/30) indicated a moderate level of confidence (“somewhat confident”), and 3% (1/30) expressed uncertainty. In the context of online communication with a representative, 70% (21/30) of the participants favored remaining anonymous before sharing contact information. Furthermore, 43% (13/30) of the participants expressed willingness to share personal information to receive personalized assistance. In terms of platform preference, most of the participants (22/30, 73%) preferred the computer browser format, 10% (3/30) preferred to use a mobile app, and 17% (5/30) did not have a preference. This preference for computer browser format directly ties into overall user satisfaction, which is a critical determinant in the net promoter score categorization. Focusing on optimizing this platform could lead to higher user satisfaction and thus more promoters.
Table 3. Intention to use, self-efficacy of use, and preferences for communication and platform.

<table>
<thead>
<tr>
<th>Questions and response items</th>
<th>Participants, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Intention to use the app</strong></td>
<td></td>
</tr>
<tr>
<td>Would you use this app?</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>27 (90)</td>
</tr>
<tr>
<td>Maybe</td>
<td>3 (10)</td>
</tr>
<tr>
<td>No</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Self-efficacy to use technology</strong></td>
<td></td>
</tr>
<tr>
<td>When it comes to your confidence in the use of technology, which of the following best describes you?</td>
<td></td>
</tr>
<tr>
<td>Uncertain</td>
<td>1 (3)</td>
</tr>
<tr>
<td>Neither certain nor uncertain</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Somewhat confident</td>
<td>17 (57)</td>
</tr>
<tr>
<td>Very confident</td>
<td>12 (40)</td>
</tr>
<tr>
<td><strong>Preferences when asking a representative</strong></td>
<td></td>
</tr>
<tr>
<td>When researching older adult services on the web, I’d appreciate the ability to ask a representative a question anonymously before sharing my contact information (0=strongly disagree; 10=strongly agree).</td>
<td></td>
</tr>
<tr>
<td>9-10 (promoter)</td>
<td>21 (70)</td>
</tr>
<tr>
<td>7-8 (passive)</td>
<td>3 (10)</td>
</tr>
<tr>
<td>0-6 (detractor)</td>
<td>6 (20)</td>
</tr>
<tr>
<td>When researching older adult services online, I would consider sharing personal information (contact information, etc) with a representative to better determine my loved one’s fit for the service (0=strongly disagree; 10=strongly agree).</td>
<td></td>
</tr>
<tr>
<td>9-10 (promoter)</td>
<td>13 (43)</td>
</tr>
<tr>
<td>7-8 (passive)</td>
<td>6 (20)</td>
</tr>
<tr>
<td>0-6 (detractor)</td>
<td>11 (36)</td>
</tr>
<tr>
<td><strong>Preference for platform</strong></td>
<td></td>
</tr>
<tr>
<td>When searching online for information on older adult care planning, would you prefer to use a computer browser or download a mobile app?</td>
<td></td>
</tr>
<tr>
<td>Computer browser</td>
<td>22 (73)</td>
</tr>
<tr>
<td>Mobile app I can download to my phone</td>
<td>3 (10)</td>
</tr>
<tr>
<td>No preference</td>
<td>5 (17)</td>
</tr>
</tbody>
</table>

**Caregivers’ Use of, and Interests in, Older Adult Care Services**

We also examined the current and past use of, as well as interests in, different types of older adult care services among the caregiver participants (Table 4). Among the participants, 63% (19/30) were currently using medical home health services, 33% (10/30) used hospice care, and 30% (9/30) employed certified financial planners. In addition, 27% (8/30) relied on house maintenance service, and 23% (7/30) engaged insurance agents and older adult law attorneys and participated in public and free older adult programs. Memory care, nonmedical home aid, and adult day care each had a 20% (6/30) use rate, while services such as assisted living, independent living, transportation, skilled nursing, medical providers, and rehabilitation facilities were used by 7% (2/30) to 13% (4/30) of the participants. Older adult living referral agents, caregiver support groups, and yard services were used by 3% (1/30) of the participants. The results suggest the diverse range of older adult care services currently being used, with medical home health being the most used service.
Table 4. Caregivers’ use of, and interests in, older adult care services (n=30).

<table>
<thead>
<tr>
<th>Older adult care services</th>
<th>Participants, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Using currently</strong></td>
<td></td>
</tr>
<tr>
<td>Home health</td>
<td>19 (63)</td>
</tr>
<tr>
<td>Hospice care</td>
<td>10 (33)</td>
</tr>
<tr>
<td>Certified financial planner</td>
<td>9 (30)</td>
</tr>
<tr>
<td>House maintenance</td>
<td>8 (27)</td>
</tr>
<tr>
<td>Insurance agent</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Older adult law attorney</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Public and free older adult programs</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Adult day care</td>
<td>6 (20)</td>
</tr>
<tr>
<td>Memory care</td>
<td>6 (20)</td>
</tr>
<tr>
<td>Professional home caregivers</td>
<td>6 (20)</td>
</tr>
<tr>
<td><strong>Have used before</strong></td>
<td></td>
</tr>
<tr>
<td>Professional home caregivers</td>
<td>13 (43)</td>
</tr>
<tr>
<td>Older adult law attorney</td>
<td>10 (33)</td>
</tr>
<tr>
<td>Rehabilitation facility</td>
<td>10 (33)</td>
</tr>
<tr>
<td>Insurance agent</td>
<td>9 (30)</td>
</tr>
<tr>
<td>Certified financial planner</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Older adult living referral agent</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Hospice care</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Transportation services</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Skilled nursing</td>
<td>6 (20)</td>
</tr>
<tr>
<td>Home health</td>
<td>6 (20)</td>
</tr>
<tr>
<td>Public and free older adult programs</td>
<td>6 (20)</td>
</tr>
<tr>
<td><strong>Would never use</strong></td>
<td></td>
</tr>
<tr>
<td>Yard services</td>
<td>8 (27)</td>
</tr>
<tr>
<td>Certified financial planner</td>
<td>7 (23)</td>
</tr>
<tr>
<td>House maintenance service</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Care manager</td>
<td>7 (23)</td>
</tr>
<tr>
<td>Insurance agent</td>
<td>5 (17)</td>
</tr>
<tr>
<td>Independent living</td>
<td>5 (17)</td>
</tr>
<tr>
<td>Assisted living</td>
<td>4 (13)</td>
</tr>
<tr>
<td>Meal service</td>
<td>4 (13)</td>
</tr>
<tr>
<td>Memory care</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Rehabilitation facility</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Older adult living referral agent</td>
<td>3 (10)</td>
</tr>
<tr>
<td><strong>Would like to learn more</strong></td>
<td></td>
</tr>
<tr>
<td>Caregiver support group</td>
<td>24 (80)</td>
</tr>
<tr>
<td>Medical providers</td>
<td>23 (77)</td>
</tr>
<tr>
<td>Memory care</td>
<td>19 (63)</td>
</tr>
<tr>
<td>Meal service</td>
<td>19 (63)</td>
</tr>
<tr>
<td>Adult day care</td>
<td>19 (63)</td>
</tr>
<tr>
<td>Older adult living referral agent</td>
<td>18 (60)</td>
</tr>
</tbody>
</table>
Regarding past use, 43% (13/30) had previously used nonmedical home aid, and 33% (10/30) had engaged older adult law attorneys and used rehabilitation facilities. Insurance agents were consulted by 30% (9/30) of the participants, and 23% (7/30) had used certified financial planners and older adult living referral agents, as well as hospice care and transportation services. Skilled nursing, medical home health, and public and free older adult programs were previously used by 20% (6/30) of the participants, while independent living and adult day care had a use rate of 17% (5/30). Memory care, assisted living, meal services, and yard services were used by 13% (4/30), while house maintenance, care managers, and medical providers were used by 7% (2/30). Caregiver support groups were attended by 3% (1/30) of the participants.

In terms of preferences for further exploration, 80% (24/30) expressed interest in learning more about care support groups, while 77% (23/30) were interested in understanding medical providers better. In addition, 63% (19/30) showed interest in memory care, meal services, and adult day care. Furthermore, 60% (18/30) were interested in older adult living referral agents, transportation services, and care managers. An additional 57% (17/30) desired to learn more about skilled nursing, assisted living, and public and free older adult programs. Interestingly, all older adult care services were mentioned as areas of interest by participants.

For services that would never be considered, some of the participants (18/30, 60%) indicated reluctance toward specific services, with yard services (8/30, 27%), certified financial planners (7/30, 23%), and house maintenance (7/30, 23%) being among those mentioned, while none expressed a definitive refusal to consider nonmedical home aid or hospice care.

### Platform Feature Evaluation Results by Participants’ Characteristics

We conducted 2-sample *t* tests to assess variances in platform feature evaluation results based on caregivers’ characteristics, providing mean scores and SDs for reporting (Multimedia Appendix 1). The evaluation results for the Olera.care platform, spanning engagement, functionality, aesthetics, information quality, and overall satisfaction, were notably consistent across all caregiver groups. However, a statistically significant difference (*P*=.02) was observed in the functionality evaluation scores, with caregivers dedicating at least 20 hours to care (mean 4.6, SD 0.4) rating it higher than those providing less care (mean 4.2, SD 0.5). In addition, caregivers with <5 years of caregiving experience reported significantly higher evaluation scores for aesthetics (mean 4.7, SD 0.4 vs mean 4.3, SD 0.7; *P*=.04) and information quality (mean 4.8, SD 0.2 vs mean 4.6, SD 0.3; *P*=.03) compared to those with a minimum of 5 years of caregiving experience.

### Discussion

#### Principal Findings

This research lays the foundation for the development of digital tools tailored to the needs of caregivers. The principal findings involve the quality and usability of the Olera.care platform, a web-based care planning tool designed to assist caregivers of people living with dementia in addressing their legal and financial needs and enable them to access functional care services. The results suggest that the Olera.care web tool is a practical, engaging, easy-to-use, visually appealing, and informative digital platform designed to provide resources that address common challenges faced by family caregivers of people living with dementia [5]. The study assessed caregivers’ intentions to use the Olera.care platform, their expectations for caregiving educational content, and their preferences for web-based information delivery. These aspects are crucial in our iterative build-measure-learn framework of research and development, which underpins our commitment to caregiver-centric product design.

The results indicate that the tested Olera.care web tool can distinguish itself not only in terms of practicality and user-friendliness but also in the quality of its content and its degree of personalization. We acknowledge existing solutions such as the Community Resource Finder by AARP and the Alzheimer’s Association, the Alzheimer’s Navigator by the Alzheimer’s Association, and CareNav by the Family Caregiver Alliance. These platforms offer valuable databases and guidance for dementia caregiving. However, our Olera.care platform differentiates itself by providing recommendations and resources that are not only categorized but also personalized and tailored to the caregiver’s specific characteristics and preferences. This unique aspect of Olera.care addresses a gap in current offerings and stands in contrast to many currently available web-based information tools such as static web pages, resource directories, or learning modules, which can be inadequate in addressing certain needs due to their lack of user engagement, personalization, relevance, and adoptability [29,30]. By contrast, the Olera.care platform attempts to involve users in the design of the platform and address these issues effectively, and the platform stands out for its interactivity, visual appeal, personalization capabilities, and informative content, making it a valuable resource for family caregivers of people living with dementia.

One noteworthy finding is that participants who devoted more weekly hours to caregiving and had limited cumulative

<table>
<thead>
<tr>
<th>Older adult care services</th>
<th>Participants, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transportation services</td>
<td>18 (60)</td>
</tr>
<tr>
<td>Care manager</td>
<td>18 (60)</td>
</tr>
<tr>
<td>Skilled nursing</td>
<td>17 (57)</td>
</tr>
<tr>
<td>Assisted living</td>
<td>17 (57)</td>
</tr>
<tr>
<td>Public and free older adult programs</td>
<td>17 (57)</td>
</tr>
</tbody>
</table>
caregiving experience tended to rate the Olera.care platform more favorably. This suggests that the platform offers specific support and benefits to caregivers with heavier caregiving workloads and those with limited prior caregiving experience. This insight underscores the importance of tailoring digital tools to the specific needs of caregivers in different situations, considering their experience and time commitment [29,30]. Such findings are integral to the build-measure-learn framework, guiding the iterative development of the platform to better align with the specific needs of caregivers. Furthermore, the relevance of the Olera.care platform is heightened in the context of the COVID-19 pandemic, which significantly disrupted family caregiving arrangements, as evidenced by more than half of these arrangements being affected [31]. This disruption led to heightened psychological burdens on caregivers, including increased depression, anxiety, and loneliness [31,32]. The pandemic also exacerbated the shortage of professional caregivers, further challenging the support systems for older adult care. Studies highlight the increased stress levels among caregivers, particularly those caring for individuals with severe dementia [33], and the overall strain on mental health resources for both caregivers and patients [34]. With the increasing social and support needs of caregivers, internet-based tools are crucial to help caregivers to access information and gain support [35].

Another significant finding pertains to the use of, and demand for, older adult care services. Many caregivers reported using or intending to use services such as home health, hospice care, insurance agents, older adult law attorneys, and financial planners. However, there is notable interest in exploring other services, such as caregiver support groups, medical providers, skilled nursing, memory care, public and free older adult programs, meal services, adult day care, and various older adult living alternatives. The data indicate that these services are often underused, potentially leading to unmet needs among older adults and their caregivers [36]. This underuse may stem from a lack of awareness about the availability of older adult care services among caregivers [37]. This insight has prompted a shift in our database curation strategy, and we are focusing now on underused yet high-demand care services and programs to make them more readily available to caregivers through our platform.

The high engagement, functionality, aesthetics, and information quality of the Olera.care digital platform can be attributed to several underlying principles and strategies. First, the platform’s unique approach of involving caregivers in the design process has proven highly effective. We demonstrated that technology interventions developed with input from the target population will increase overall satisfaction with, and preference for, the product. Second, the curated content and resources of the Olera.care platform, informed by leading experts, have been rated highly relevant by the participants, with 73% (22/30) reporting the content as highly relevant with a mean MARS score of 4.76 (SD 0.44) of 5. This demonstrates the significance of expert guidance in creating a resource that resonates with the target audience [26]. The curated content and resources on the Olera.care platform can be further improved through the incorporation of artificial intelligence and large language models, allowing for improved and personalized recommendations based on an existing recommendation system [38]. Large language models can process large data sets across numerous relevant variables (eg, specific needs, geographic location, and financial constraints) to provide the most appropriate care solutions. To increase the accuracy of the large language model, the data would have to undergo rigorous quality control and standardization. Industry experts that we are currently working with to inform our curated content could also validate the model through feedback on its accuracy and quality. The model’s recommendations can also be consistently improved through the input of new data, further increasing the accuracy of its recommendations.

Limitations and Future Research and Practice

It is important to acknowledge the limitations of this pilot study, which we view as opportunities for further learning and refinement within our build-measure-learn framework. First, the small sample size and lack of racial and ethnic diversity among the participants may limit the generalizability of the findings, pushing us to expand our research scope. The participants in this study were also mostly technology savvy and well educated. However, the demographic characteristics of the caregiver participants in this study aligned well with the caregiver profile in the United States (ie, the majority were women: 23/30, 77%; aged ≥50 years: 25/30, 83%; and non-Hispanic White: 25/30, 83%) [39]. The study populations also represented some diversity in financial levels and caregiving experiences. Future research should aim to include a more racially and ethnically diverse study population to ensure a broader representation of caregiver experiences and preferences. Second, all acceptability and usability metrics in this study were self-reported, which could introduce self-report bias. However, we used a validated tool, the MARS [25], which was adapted according to platform features. Third, given the limited sample size and insufficient statistical power in the subgroup analyses, both significant and nonsignificant t test results should be interpreted cautiously. Furthermore, the study’s short duration of interaction with the digital platform also limits our understanding of the platform’s long-term usability and usefulness for caregivers of people living with dementia.

A future study should be conducted with a larger and more diverse group of caregivers, allowing for a more comprehensive assessment of the platform’s long-term perceived usability and ease of use, thus enabling us to continually learn and refine the platform. For future practice, we are focused on enhancing the accessibility and visibility of our web application for family caregivers by integrating it into the existing care delivery framework through strategic digital marketing, primarily using organic search engine optimization. This strategy ensures that our platform aligns with user search behaviors, making it easily discoverable by those in need of caregiving resources. Simultaneously, we are committed to maintaining universal accessibility, opting for a broader reach via effective search engine optimization strategies over direct integration into health plans. Our platform, designed more as a comprehensive digital health tool than a conventional medical device, provides holistic care planning support for family caregivers. Financially, we have chosen a sustainable business-to-business revenue model, focusing on advertising and commission-based referral fees.
from businesses serving our caregiver community, which allows us to offer our services free of charge to families. This approach is underpinned by the older adult care industry’s potential for low-cost client acquisition, enabling us to provide much-needed support to families without financial burden and ensuring that our platform remains accessible to all caregivers in need.

Conclusions
The Olera.care platform, characterized by its practicality, interactivity, ease of use, visual appeal, and informativeness, shows promise as a valuable tool for dementia caregivers. With the pilot group of caregivers’ engagement and feedback, the platform provides tailored support to meet the specific challenges of dementia caregiving. Future development and research are essential to enhance the platform and comprehensively evaluate its efficacy in supporting caregivers and alleviating caregiving burdens across broader and more diverse populations.

Acknowledgments
The authors acknowledge the support provided by the Center for Community Health and Aging at Texas A&M University and their community partners. The authors extend their sincere appreciation to the family caregiver participants’ engagement and contributions to evaluating the Olera.care platform. This study was supported by the National Institute on Aging Small Business Innovation Research program (1R44AG074116-01).

Authors’ Contributions
TF, LD, SL, and MGO contributed to study design. LD, TF, MNH, JV, and QF contributed to study implementation and data collection. QF contributed to data analysis. QF, MNH, MGO, and LD contributed to data interpretation. SL and MGO contributed to supervision. QF, MNH, and LD contributed to the original draft writing. All authors contributed to critical revision of the manuscript.

Conflicts of Interest
LD is an executive and owner of Olera, Inc.

Multimedia Appendix 1
Results of the 2-sample 2-tailed t tests for platform feature evaluation by participants’ characteristics.
[DOCX File, 377 KB - aging_v7i1e55132_app1.docx ]

References


27. Mandracchia F, Llauradó E, Tarro L, Valls RM, Solà R. Mobile phone apps for food allergies or intolerances in app stores: systematic search and quality assessment using the mobile app rating scale (MARS). JMH Q Health U Health 2020 Sep 16;8(9):e18339 [FREE Full text] [doi: 10.2196/18339] [Medline: 32936078]


Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>Alzheimer disease</td>
</tr>
<tr>
<td>MARS</td>
<td>Mobile Application Rating Scale</td>
</tr>
</tbody>
</table>

©Qiping Fan, Minh-Nguyet Hoang, Logan DuBose, Marcia G Ory, Jeswin Vennatt, Diana Salha, Shinduk Lee, Tokunbo Falohun. Originally published in JMIR Aging (https://aging.jmir.org). 17.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
The Use of Digital Technologies in the Promotion of Health Literacy and Empowerment of Informal Caregivers: Scoping Review

Suzete Soares1,2*, MSci; Louíse Viecili Hoffmeister3,4*, MSc; Maria de Fátima Fernandes1,5*, MSc; Adriana Henriques3,6*, PhD; Andreia Costa3,6*, PhD

1Nursing School of Lisbon, Lisbon, Portugal
2Family Health Unit Carnide Quer, North Lisbon Health Center Cluster, Lisbon, Portugal
3Nursing Research, Innovation and Development Centre of Lisbon (CIDNUR), Nursing School of Lisbon, Lisbon, Portugal
4National School of Public Health, NOVA University, Lisbon, Portugal
5Community Care Unit Integrar na Saúde, Administração Regional de Saúde de Lisboa e Vale do Tejo, Lisbon, Portugal
6Institute of Environmental Health, Faculty of Medicine of the University of Lisbon, Lisbon, Portugal

*all authors contributed equally

Abstract

Background: Informal caregivers (IC) play an important role in the community as health care providers for people who are dependent on self-care. Health literacy contributes to empowerment, better care, and self-management of one’s own health and can be developed using digital technologies.

Objective: This study aims to map scientific evidence about the use of digital technologies to promote health literacy and the empowerment of ICs.

Methods: We conducted a scoping review following the Joanna Briggs Institute methodology. The CINAHL, MEDLINE, Scopus, and PubMed databases were searched to find primary studies on the theme. Inclusion criteria were based on the Population, Concept, and Context logic. To be selected for analysis, studies must have involved informal or family caregivers aged ≥18 years who provide care to dependent persons and who have access to the internet and digital devices (computer, smartphone, and tablet). A total of 2 independent researchers (SS and LVH) performed the screening process. This study is part of a main project that was approved by the Ethics Committee for Health of the Regional Health Administration of Lisbon and Tagus Valley (reference 058/CES/INV/2022).

Results: A total of 9 studies were included in the review. The analysis of the studies showed that ICs use digital tools, such as computers and smartphones, with smartphones being the preferred tool. ICs use the internet to access information; manage home tasks; communicate with relatives, their peers, and health care professionals; and take part in forums. Due to difficulties in leaving their houses, forums are highly valued to preserve human connections.

Conclusions: The use of digital technologies to convey clear, objective, reliable, and accessible information is a strategic action for promoting health literacy and for contemplating the variable care needs of ICs. By working with ICs in the development of new technologies, researchers are building a new tool that meets ICs’ needs.

(JMIR Aging 2024;7:e54913) doi:10.2196/54913

KEYWORDS

health literacy; empowerment; digital technology; informal caregiver; family caregiver
Introduction

Background

Population aging is a social challenge worldwide. As life expectancy increases, the incidence of chronic and incapacitating diseases also increases [1]. The high levels of dependence and the complex management of health status raise awareness of the increasingly relevant role of informal caregivers (ICs) in terms of care and health promotion of their relatives [2]. The ICs, defined as someone who provides nonremunerated care to a person with a long-term illness, disability or other health need, or long-term care, outside a professional or formal framework, are considered essential both to the care of people in the community and to the economy of European Union countries [3]. ICs are more and more important to patients as well as to health care professionals [4]. They play a central role in the planning, training, and provision of services to people with care needs [2,3]. In practical terms, ICs collaborate in providing health care at home to people who show an impairment in self-care, activities of daily living, and instrumental activities of daily living. The needs change over time, as does the level of dependency of the person cared for [5].

In most situations, ICs are not prepared to play this role. This transition in caregivers’ lives brings out feelings of insecurity due to the unknown and the lack of knowledge and skills to ensure that the person cared for is given proper care [5]. This way, ICs’ existing and acquired knowledge is extremely important and becomes necessary for the implementation of interventions destined to promote the development of skills and the involvement of relatives in patient trajectories to improve patient outcomes [2-4].

A vital point in health policies is the investment in the health literacy (HL) of ICs. HL is central to empowering people, their families, and communities, promoting greater control over decisions and actions affecting their health [6]. HL is defined as the ability to access, understand, evaluate, and apply information about health care, disease prevention, and health promotion to maintain and promote quality of life during the life course [7].

Through HL development, conditions are created for individuals to gain knowledge and skills, make informed decisions, and feel motivated to adopt a behavior that improves their health status and well-being [8].

Considering that HL is a health determinant, mediator, and moderator, it is important to ensure that citizens access reliable, useful, and updated health information to help them make the best decisions about their personal health, their family’s health, and the community’s health [9-11]. Proper access to information allows to promote and increase citizen empowerment so that they participate in their health care, leading to shared responsibility and informed decision-making [12].

As an agent, the health care professional plays a central role in effective communication and in conveying reliable information to the population. User-relative–health care professional communication significantly affects health outcomes and user satisfaction concerning health services [13]. Digital technologies have created an opportunity for health professionals and health organizations to directly communicate with many people in real time. This digital revolution in communication allows to customize information, help people set health targets, and interact in real time [1].

Information and Communication Technologies (ICTs) is the set of technologies and equipment that, in an integrated manner, allow working and communicating information, including computers and the respective applications, the internet, and telecommunications [14]. They are part of the citizens’ routine, with an increasing use of educational platforms. The internet is considered a privileged means of interaction with the population that needs health care [1]. ICTs improve the quality of life of older adults and their caregivers and their access to quality care, contributing to improving the social lives of caregivers and decreasing their isolation via social activities and intergenerational relationships [15]. These aspects contribute to balanced physical, mental, and emotional health and to a decrease in depressive symptoms and sadness. Digital technologies are considered a key component and facilitator of sustainable health systems and universal health coverage [16]. Digital technology is a strategy that can promote accessibility to health care for all citizens. Digital means can be used to increase access to reliable, useful information and to strategies that meet the needs of the highest possible number of ICs, whether in real time or not [1]. However, accessing and handling these technologies requires digital literacy, which is one of the barriers identified by studies in certain groups considered vulnerable, such as older adults. Digital HL is the ability of citizens to use digital platforms to manage their health, validate web-accessible health information, and communicate with health professionals [17].

Objectives

In Portuguese literature, there are only a few scientific studies conducted by nurses that refer to the use of digital technologies as a resource to empower dependent people and family caregivers [1]. The need to know if dependent people and their ICs have access to digital technologies and use them when they have health needs gave rise to the following research question: “Which digital technologies are used for promoting Health Literacy and empowering the Informal Caregiver?” For the mapping, we used the following guiding questions: “Do the Informal Caregivers have access to digital technology?” and “Do the Informal Caregivers use digital technologies to improve their health literacy and empowerment concerning the care of the person cared for?” To answer these questions, this review aims to map the scientific evidence regarding the use of digital technologies to promote HL and empower ICs.

Methods

Overview

This is a scoping review conducted according to the methodology recommended by the Joanna Briggs Institute (JBI) [18]. Scoping reviews are used to identify knowledge gaps,
enhance knowledge described in the literature, clarify concepts, or investigate research conduct [19].

The theme was searched in the JBI Database of Systematic Reviews, CINAHL, MEDLINE, Scopus, and PubMed, and no systematic review was found for this same theme. The inclusion criteria were based on the Population, Concept, and Context logic: the Population included all informal or family caregivers aged ≥18 years who provide care to dependent persons and who have access to the internet and digital devices (computer, smartphone, and tablet).

The search was conducted from April 4 to 18, 2022, and included primary qualitative and quantitative studies and mixed method studies in English, Portuguese, French, and Spanish, during 5-year period between January 2017 and December 2021, to obtain the most recent studies published on this theme. Key terms and inclusion criteria were used as a strategy to identify papers that were relevant to the search.

**Study Selection, Data Extraction, and Analysis**

According to the JBI’s recommendations, the search strategy was performed in 2 steps [18]. There was an initial search of the electronic platform EBSCO, in particular, MEDLINE and CINAHL, with the natural keywords informal caregiver; family caregiver; health literacy; empower; digital technology; and community, following the search for the indexing term MH “Empowerment.” Subsequently, we carried out an analysis of the words used in the title, the abstract, and the terms indexed as well as the keywords presented in the description of each searched article. We then carried out a second survey in which the indexing terms and keywords were searched in MEDLINE (PubMed), CINAHL (via EBSCO), MEDLINE (via EBSCO), and Scopus (Textbox 1).

A total of 2 independent reviewers (SS and LVH) analyzed the relevance of papers using the information included in the title and abstract, considering that the study population must be defined and the goal must be associated with digital tools.

It was necessary to retrieve the papers after reading the abstract. Full papers were obtained for all studies with the inclusion criteria. A table was filled with the defined criteria, considering the goals and the results of the study that would answer the research question. After reading the full text, 2 papers showed a divergent opinion. This situation was discussed and resolved without the need to speak to a third reviewer.

The screening process identified 442 studies. Of the 442 studies, 77 (17.4%) were duplicated and so were excluded. Of the remaining 365 studies, 320 (87.7%) were excluded for their titles and 25 (6.8%) for their abstracts, based on the inclusion criteria that had determined their eligibility. In the second step, there were 20 papers for full-text review. Of the 20 papers, 11 (55%) were eliminated because of the following reasons: 4 (20%) because their goals did not relate to the technological needs of ICs but to the person cared for; 4 (20%) because they were about another type of nontechnological experience; and 3 (15%) because they were about behavioral therapies and coping strategies. Figure 1 [20] shows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart for the identification and selection of the studies.

A data collection instrument was made to extract information from the selected studies, including the following items: author, country, year of publication, study goal, study type and methodology used, population, sample, types of interventions, main results, and conclusions. The results were analyzed based on their content and organized according to the research question and goals.
Textbox 1. Search strategy according to database searched.

**MEDLINE (via PubMed)**
- (((informal caregivers) OR (family caregivers) AND (community) AND ("health literacy") OR (empowerment) OR ((digital education) OR (digital technology) OR (digital era) OR (digital platforms) OR (digital sources) OR (Information and communication technology))) in the last 5 years)

**CINAHL complete (via EBSCO)**
- S1 informal caregivers
- S2 family caregivers
- S3 S1 OR S2
- S4 community
- S5 health literacy
- S6 empowerment
- S7 MH"empowerment"
- S8 empower*
- S9 digital technology
- S10 digital era
- S11 digital health literacy
- S12 digital sources
- S13 digital education
- S14 digital platforms
- S15 Information Communication Technology
- S16 S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15
- S17 S3 AND S4 AND S16

**MEDLINE (via EBSCO)**
- S1 informal caregivers
- S2 family caregivers
- S3 S1 OR S2
- S4 community
- S5 health literacy
- S6 empowerment
- S7 MH"empowerment"
- S8 empower*
- S9 digital technology
- S10 digital era
- S11 digital health literacy
- S12 digital sources
- S13 digital education
- S14 digital platforms
- S15 Information Communication Technology
- S16 S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15
- S17 S3 AND S4 AND S16

**Scopus**
-
Ethical Considerations
This study is part of a main project that was approved by the Ethics Committee for Health of the Regional Health Administration of Lisbon and Tagus Valley (reference 058/CES/INV/2022).

Results

Studies Characteristics
In total, 9 studies were included in this review after the search. According to the JBI’s guidelines, after the selection, the studies were organized based on their research question and goals (Tables 1 and 2).
Table 1. Summary of study characteristics.

<table>
<thead>
<tr>
<th>Paper title</th>
<th>Authors, country, and year</th>
<th>Goal</th>
<th>Type of study and methodology</th>
<th>Population and sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is “care quality” and can it be improved by information and communication technology? A typology of family caregivers’ perspective</td>
<td>Leslie et al [21], Canada, 2020</td>
<td>• To determine how ICTs can support family caregivers who play the caregiver role</td>
<td>Qualitative study Method</td>
<td>FC of older people</td>
</tr>
<tr>
<td>Effect of an innovative model of complexity care on family caregiver experience: qualitative study in family practice</td>
<td>Nickell et al [22], Canada, 2020</td>
<td>• To learn about the experiences of FC of older people with complex needs, using the Interprofessional Model of Practice for Aging and Complex Treatments</td>
<td>Qualitative study Method</td>
<td>FC of older people with complex needs</td>
</tr>
<tr>
<td>Building a Research Roadmap for Caregivers Innovation: Finding from a Multi-Stakeholder Consultation and Evaluation</td>
<td>Egan et al [23], Scotland, 2021</td>
<td>• To explore a future roadmap for innovation from IC participation</td>
<td>Mixed study Method</td>
<td>ICs</td>
</tr>
<tr>
<td>The care capacity goals of family carers and the role of technology in achieving them</td>
<td>Leslie et al [24], Canada, 2020</td>
<td>• To identify the goals of FC when caring for older adults and how technology can help achieve those goals</td>
<td>Mixed study Method</td>
<td>FCs</td>
</tr>
<tr>
<td>Mobile Support for Older Adults and Their Caregivers: Dyad Usability Study</td>
<td>Quinn et al [25], United States, 2019</td>
<td>• To determine the usability of a mobile app within the older population and in their relationship with ICs</td>
<td>Observational study Method</td>
<td>Older people and ICs</td>
</tr>
<tr>
<td>Preferences for using a Mobile App in Sickle cell Disease Self-management: descriptive Qualitative study</td>
<td>Mayo-Gamble et al [26], Canada, 2020</td>
<td>• To explore health preferences for using an app in the process of facilitating the self-management of adults with sickle cell disease and their caregivers who live in urban and rural communities</td>
<td>Qualitative study Method</td>
<td>Adults with sickle cell disease and caregivers</td>
</tr>
<tr>
<td>A Digital Mobile Community App for Caregivers in Singapore: predevelopment and Usability Study</td>
<td>Lwin et al [27], Singapore, 2021</td>
<td>• To provide a clear understanding of the implementation along with a usability study to gauge user opinion of the “Caregiver’s circle” app within Singapore</td>
<td>Qualitative study Method</td>
<td>ICs</td>
</tr>
<tr>
<td>Improving the Quality of Life of Family Caregivers of People with Alzheimer’s Disease through Virtual Communities of Practice: A Quasiexperimental Study</td>
<td>Romero-Mas et al [28], Spain, 2021</td>
<td>• To describe the relation between the quality of life of ICs of people with Alzheimer disease and their participation in a VCoPs (virtual community with the exchange of knowledge and an emotional support and collaboration culture)</td>
<td>Quasiexperimental study Method</td>
<td>ICs of people with Alzheimer disease</td>
</tr>
<tr>
<td>Patient Portals as a Tool for health Care Engagement: A Mixed-Method Study of older Adults with Varying Levels of Health Literacy and Prior Patient Portal Use</td>
<td>Irizarry et al [29], United States, 2017</td>
<td>• To explore attitudes in relation to choosing the portal and its utility as a tool to involve health care with different levels of HL</td>
<td>Mixed-methods study method</td>
<td>Older people</td>
</tr>
</tbody>
</table>

*ICs* = Family carers; *FCs* = Family care coordinators; *FC* = Family carers; *IC* = Family carers.

Soares et al
aICT: Information and Communication Technology.
bFC: family caregiver.
cIC: informal caregiver.
dVCoP: virtual community of practice.
eHL: health literacy.
Table 2. Summary of study results.

<table>
<thead>
<tr>
<th>Types of interventions</th>
<th>Main findings</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Intervention made with 10 focus groups from May 2017 to August 2018. Each session took 2 hours.</td>
<td>Technologies that are only focused on the task can lose their value as they lose the capacity to provide information that is relevant to caregivers’ needs.</td>
<td>ITC product development supported by ICs should focus on human relationships and expand a facilitating communication, allowing their participation in decision-making and allowing them to express their concerns and goals.</td>
</tr>
<tr>
<td>• Bottom-up approach with thematic content analysis.</td>
<td>ICTs, as the intermediary for an improvement in quality of life and as providers of relevant information that are enabled with knowledge and caregiver needs’ change.</td>
<td>Technology appears as a support to receive information that is relevant to caregivers’ needs and to establish human connections.</td>
</tr>
<tr>
<td>• Individual interviews with 13 family caregivers about the caregiver role and their (physical and emotional) well-being.</td>
<td>Caregivers reported that they no longer felt lonely in this role because they were given basic information about the disease as well as existing resources and equipment. They felt recognized and heard; they were able to express their uncertainties, stories, and suggestions, increasing their commitment to caregiving. They searched the internet.</td>
<td>Involving ICs as part of the multiprofessional team increases their perception and understanding of the caregiver role and their trust in their ability to perform this role and facilitates their empowerment.</td>
</tr>
<tr>
<td>• The patient and the caregiver are encouraged to play a more active role in the process of their disease by raising questions and discussing actions.</td>
<td>In total, 108 of the 112 (96%) ICs use digital technology.</td>
<td>A technological approach in the following areas is required for the health and well-being of ICs: information, monitoring technology, and communication with other ICs and professionals.</td>
</tr>
<tr>
<td>• A mixed approach was used:</td>
<td>In total, 108 of the 112 (96%) ICs use digital technology.</td>
<td>The experience of ICs in collaborating with universities to identify priorities and actions that speed up searches and future political decisions about significant and innovative solutions should be valued.</td>
</tr>
<tr>
<td>• A 10-minute multisector consultation from June 15, 2020, to September 30, 2020.</td>
<td>The hybrid approach (both in person and web based) can work for caregivers. The experience of ICs in collaborating with universities to work via multiple communication channels should be valued. Deep knowledge of needs and existing gaps allows one to contribute to technological innovation to overcome existing technological barriers and learn what the facilitating mechanisms are. The ICs mentioned the need for improved financial, emotional, psychological, training, and educational support.</td>
<td></td>
</tr>
<tr>
<td>• Web-based questionnaire on social media.</td>
<td>In total, 108 of the 112 (96%) ICs use digital technology.</td>
<td></td>
</tr>
<tr>
<td>• Sequential method, focus group, and web-based questionnaire. In total, 10 focus groups with 25 family caregivers. The intervention took place from May 2017 to August 2018.</td>
<td>Technology maintains the ability to care and allows ICs to develop coping strategies, guide themselves, and socialize. Technology is an intermediary that connects ICs to information support and other caregivers. Key targets for ICs are to reinforce and preserve their ability to provide care.</td>
<td>Technology is well positioned to find the best self-care to facilitate the connections needed for a social life.</td>
</tr>
<tr>
<td>• First part: discussion of targets and technological solutions. In what they think technology can help them. Second part: web-based questionnaire about 7 fields: physical health, mental health, well-being, social connection, education, employment, and finances.</td>
<td>Technology maintains the ability to care and allows ICs to develop coping strategies, guide themselves, and socialize. Technology is an intermediary that connects ICs to information support and other caregivers. Key targets for ICs are to reinforce and preserve their ability to provide care.</td>
<td>Technology is well positioned to find the best self-care to facilitate the connections needed for a social life.</td>
</tr>
<tr>
<td>• Technological targets and suggestions should imply that the understanding of care as a source of overload was transformed into a more resilient, sustainable caregiving model.</td>
<td>Technology is well positioned to find the best self-care to facilitate the connections needed for a social life.</td>
<td></td>
</tr>
<tr>
<td>• Technology can help promote such resilience but can be limited to the role of an intermediary that connects family caregivers to information supports and peers.</td>
<td>Technology is well positioned to find the best self-care to facilitate the connections needed for a social life.</td>
<td></td>
</tr>
</tbody>
</table>

https://aging.jmir.org/2024/1/e54913

JMIR Aging 2024 | vol. 7 | e54913 | p.324

(page number not for citation purposes)
### Types of interventions

- Participants completed a skill evaluation questionnaire and downloaded an app to their smartphones or computers that was used for a month.
- Then, participants completed 2 questionnaires that evaluated app features and aesthetics and their relationship with the app.
- App features: user profile, family health history, health information, receiving studies based on their health profile, and establishing a relationship with their caregivers.

### Main findings

- Study results showed normal levels of digital competence for the older adults and high levels for the ICs. Older adults use their smartphones to make calls (9/12, 75%) and read emails (7/12, 58%). They access the internet (4/12, 33%) but on their computers. ICs use their smartphones for calls, SMS text messages, emails, and the internet equally (11/12, 92%). They access the internet via their phones.
- This study concluded that 50% (6/12) of ICs want to use the app to manage the appointments and clinical information of the person cared for and to access specific information that allows them to share and discuss to commit to the caregiving; they believe that the app’s esthetic dimension is important.

### Conclusion

- Technologically experienced caregivers play an essential role in showing the benefits of technology for supporting care provision for older adults.
- There were high levels of use of technology among the older adults and caregivers, but there was only an average use of the mobile app. Additional training is recommended for the older adults and caregivers, including behaviors directed toward keeping digital health records.

<table>
<thead>
<tr>
<th>Types of interventions</th>
<th>Main findings</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Participants completed a skill evaluation questionnaire and downloaded an app to their smartphones or computers that was used for a month.</td>
<td>Study results showed normal levels of digital competence for the older adults and high levels for the ICs. Older adults use their smartphones to make calls (9/12, 75%) and read emails (7/12, 58%). They access the internet (4/12, 33%) but on their computers. ICs use their smartphones for calls, SMS text messages, emails, and the internet equally (11/12, 92%). They access the internet via their phones.</td>
<td>Technologically experienced caregivers play an essential role in showing the benefits of technology for supporting care provision for older adults.</td>
</tr>
<tr>
<td>- Then, participants completed 2 questionnaires that evaluated app features and aesthetics and their relationship with the app.</td>
<td>This study concluded that 50% (6/12) of ICs want to use the app to manage the appointments and clinical information of the person cared for and to access specific information that allows them to share and discuss to commit to the caregiving; they believe that the app’s esthetic dimension is important.</td>
<td></td>
</tr>
<tr>
<td>- App features: user profile, family health history, health information, receiving studies based on their health profile, and establishing a relationship with their caregivers.</td>
<td>Participants are receptive to using the app to self-manage the disease. A mobile app reduces the information access barrier. In rural communities, the app increases ICs’ access to resources.</td>
<td>The results can be used to develop a patient-centered health app that is easy to use to facilitate disease self-management, thereby increasing access to resources by relatives that live in rural communities.</td>
</tr>
</tbody>
</table>

- In total, 5 community listening sessions were made with 1 urban and 1 rural community. Each session took 2 hours. A questionnaire about demographics and access to technology was applied.
- Where they searched for information about self-care in relation to the SCD and what was their satisfaction level with the search for and support about management and resources. A total of 7 aspects were evaluated: self-management information, such as receiving information, which information they wish to receive, changes in disease management, support types, barriers to and facilitators for the use of apps, and mobile app preferences.

- A predevelopment survey was made about the following issues: care, support provided, and what they would like in a caregiving mobile app. Identifying the needs of ICs and the gaps in web community networks.
- Demographics about the health of the person cared for and about the ICs’ physical and mental health.
- What is the level of use of digital means when searching for information and support.
- A total of 32 caregivers completed a web-based questionnaire and in-person interviews, followed by a usability test.

- Caregivers enjoyed the “Caregivers’ Circle” and were confident that this app could help them improve their quality of life.
- Including many resources that caregivers need daily in 1 app can help save time and help them live without problems.
### Types of interventions
- ICs said they liked using the app. They said it was useful, easy to use, and helpful to improve the quality of life because they included multiple resources: a public forum for discussions with the community and other ICs in the same region without ever leaving home and a market to purchase and sell material and equipment required for caregiving. Including many resources that caregivers need daily in an easy-to-use app allowed them to save time and helped browse without any issues.
- The use of smartphones created an opportunity for the caregiving community to use technology in a useful way.
- The app included caregivers’ ideas, which created an app that facilitated caregiving.
- As to concerns about safety and security, trust would increase if the app were supported by a renowned organization.
- ICs have suggested that the app should include a resource that would help with mental health, namely, relaxation techniques, motivational quotes, and guides that would remind them to take care of themselves.

### Main findings
- QoL was 66.6 and increased to 69.5.
- There was no discrepancy between sexes for the QoL. Age was the only sociodemographic criterion that affected the quality of life; older adults increased their QoL to 74.6. Young people went from 66.7 to 67.85. Spouses said that the app had a positive impact on their QoL.
- Regarding HL, the average rate of 26.10 (in 40) increased to 30.68. Internet interventions can help caregivers meet their needs, which is a positive experience.
- Allowed to get to know their peers and to feel less lonely.

### Conclusion
- Caregivers can benefit from the VCoP because it enables interaction and knowledge sharing between caregivers and helps them meet their needs.
- VCoP’s impact is governed by age and relationship with the person cared for. It was positive for the caregivers’ quality of life, at a physical level, when the functional condition of the person with Alzheimer disease worsened.
- The VCoP was considered a useful tool.
- HL had a positive impact on the physical area of the QoL of caregivers.

### Methods
- The study took place between July 2017 and April 2018.
- Previous contact was made with the AF-MADO association, and explanatory sessions were held (individual and group). In total, 2 groups were created, 1 with and 1 without health care professionals. Intervention: developing an app based on the CoP theory, with space for chatting and a member file with information about each member.
- The following aspects were evaluated before and after the VCoP intervention: quality of life, HL, and the Barthel scale associated with the Spanish population.
- First contact made by phone (data collection: demographics, health, “Deficit of quality-of-life technology” questionnaire, and CREATE).
- Classified participants according to their HL level and portal use. This classification resulted in 4 groups (group 1: high HL, yes portal; group 2: high HL, no portal; group 3: low HL, yes portal; and group 4: low HL, no portal).
- Second contact made with 4 focus groups (N=75) aimed at analyzing participants' attitudes. Sessions took 1 hour, were recorded, and used NVS.
- Second contact made with 4 focus groups (N=75) aimed at analyzing participants' attitudes. Sessions took 1 hour, were recorded, and used NVS.
### Main Findings

Table 2 shows the main findings and conclusions of the studies described in the papers.

Regarding the year of publication, the studies were published in year 2017 (1/9, 11%); year 2019 (1/9, 11%); year 2020 (4/9, 44%); and year 2021 (3/9, 33%). They were conducted in the following countries: Canada (4/9, 44%), the United States (2/9, 22%), Scotland (1/9, 11%), Singapore (1/9, 11%), and Spain (1/9, 11%). Of the 9 studies, 3 (33%) followed a qualitative approach, 4 (44%) followed a mixed approach, 1 (11%) was observational, and 1 (11%) was a quasi-experiment.

The results of the studies enabled us to address the guiding questions. On the question “Does the Informal Caregiver have access to digital technology?” the studies show that ICs have access to and use digital technology [23,25]. They describe which types of technological resources are used more frequently by them: smartphones with mobile apps or internet access. The internet is the source of choice for accessing health information and learning about self-management techniques, with the importance of reliable websites being emphasized [22,25,26]. Smartphones are used to make calls, send SMS text messages and emails, and access the internet [25]. Apps are used to manage the appointments and medical information of the person cared for and to access specific information that allows ICs to share and discuss to commit to the caregiving relationship [25]. The esthetic dimension, ease of configuration, and nice interface are app features that are valued by ICs [25,26].

Privacy and security issues seem to be a factor that limits the use of technologies because users feel insecure due to the risk of sharing personal information [19,27]. Struggling to identify relevant and reliable information is also a factor that causes apprehension when it comes to internet use [26]. The degree of trust when accessing digital information seems to be related to the HL level of users [29]. Users with low HL levels who did not use the technology that was being analyzed had little experience using computers, no training, and no internet access at home. Those who used the portal showed increased interest in learning and practicing with the new technologies [29].

### Types of interventions | Main findings | Conclusion
---|---|---
- Participants with the higher HL who use the portal struggle to solve issues without the digital support and feel more pressured to use these methods. - Those who do not use the portal say they do not feel safe using it due to the risk of sharing personal data and prefer to use the phone. People with low HL who do not use the portal do not have experience using computers, are not trained, and do not have internet access at their homes, but those who use the portal say they are more interested in learning and training with new technologies. - People who are more familiar with accessing health information using the internet might be more willing to participate in research related to digital technology. The study revealed that HL was a factor that contributed to trust when accessing digital health information. However, it was not directly related to the motivation to get involved in health care. If portal users understand the benefits, this would be a motivation for portal use. Specific technology training is required to gain trust. ICs play a potential role in improving access to portal use for older adults who cannot access portals.
fact reinforces the need for specific training on the use of digital technologies to gain trust [29].

Regarding the question “Does the Informal Caregiver use digital technology to improve their HL and their training in caring for the person cared for?” the studies show that the use of digital technology can benefit the population as well as caregivers [21,25-27,29].

ICTs lower information access barriers and provide relevant information that is enabled when there is a need to gain new knowledge [21,26]. ICTs are perceived as giving ICs the opportunity to guide themselves and interact with other caregivers, which allows them to get to know their peers and feel less lonely [24,28].

Caregivers say that by using digital technologies, they can obtain basic information about the disease, such as symptoms and treatment options, and about existing resources and equipment, which makes them feel less lonely in this role [21-23]. With the support of digital tools, caregivers felt recognized and heard and could express their uncertainties, stories, and suggestions, which increased their commitment to care provision [22]. Technology also maintains their caring ability and allows them to develop coping strategies [24].

The use of technology is also referred to as an intermediary for an improved quality of life [21,28]. This perception of the improvement of the quality of life is boosted when the technology that is used includes multiple resources, such as the fact that there is a public forum for community discussion with other ICs in the same region without having to leave home, a market to purchase and sell materials and equipment that is needed for providing care, and an alert system or information trackers [26,27]. The integration of the multiple resources that are needed by caregivers daily in an app that is easy to use allows them to save time and provide help to browse without problems [27].

Another aspect referred to by the studies concerns suggestions or factors that can improve the experience of ICs when using digital technologies. One study described that new technologies that are only focused on the task can lose their value as they lose the capacity to provide information that is relevant to caregivers’ needs [21]. It is important that ICs collaborate in the development of technologies because their deep knowledge of the needs and existing gaps contribute to technological innovation, which allows them to overcome the existing technological barriers and learn facilitator mechanisms [23].

The expectations of ICs as to digital technologies also seemed to be an important aspect to consider because they can increase the technology used. ICs hope that technologies can provide emotional and psychological support, informative support from the family, training and education, and health care follow-up [23,26]. In a more practical way, ICs suggested that there should be resources that help them with their mental health, namely, relaxation techniques and motivational quotes and guides that would help them remember to take care of themselves [27].

Discussion

Informal Caregivers’ Role in the Health Care System

According to the studies that were analyzed, demographic changes are leading to an increasing need for long-term care, which results in people informally caring for their relatives. Being an informal or family caregiver brings uncertainties, isolation, and overload [21,27]. Studies have shown that the involvement of the caregiver in the care plan is essential. The active involvement of ICs as a member of the interprofessional care team results in an improved experience, increased caregiver capacity, and the appreciation of the caregiver role [22,24].

These results are in accordance with the literature where ICs are considered “one of the elements of the sustainability of social and health systems” [30]. This emphasizes how important it is for health care professionals to work with ICs to find the strategies that are most adequate for effective empowerment [30]. The empowerment of ICs should be “a priority in health care organizations and the nurse assumes a major, dynamic, empowering role when it comes to the most adequate response to meet those needs” [31].

The World Health Organization (WHO) has defined a long-term strategy for the expansion and use of digital health, emphasizing the positive impact that it can have on health care access and provision as well as on the health and well-being of the population and caregivers [23]. According to the literature, health technology is “one of the strategies used by the health care professional to empower citizens to use it in a secure way” [32].

The Use of Digital Technologies Supporting Caregivers

The studies revealed that low HL was a barrier to accessing digital information and the correct use of technological tools. Lack of training makes browsing difficult and results in user insecurity [26,29]. The initial findings of a European survey on population HL carried out by the WHO Action Network on Measuring Population and Organizational Health Literacy indicate that 22% to 58% of the population find it challenging to access and interpret digital health information [33]. By contrast, the European data report shows that, in 2019, in European countries such as Finland, the Netherlands, the United Kingdom, and Germany, 75% of the active population had basic digital skills [34].

Promoting HL improves safety in caregiving and decreases the risk associated with this activity [32].

Using digital technology in the health field can benefit the caregivers and the general population [21,25-27,29].

Questions about privacy and security when using these digital tools are an important factor for users. Although there is an increasing concern about what is the best way to develop emerging web-based technologies (eg, ethical data use), the results show that a hybrid model with a web-based and in-person approach can work well for caregivers in rural areas [23]. The model that includes digital technology and an in-person approach is pointed out as a more reliable model for the ICs.
These results are in line with the American study that described that ICs use the internet (77.5%) to access health information for themselves (73%), for others (67.5%), and to communicate with the physician [35].

The results highlight that ICs intend to use digital tools to establish communication relationships with people cared for, their family members, the peers, and health care professionals [21,23,24,27,28].

**Principal Findings**

The text highlights privacy concerns limiting technology use, underscoring the impact of low HL on users’ digital engagement. ICs benefit from digital tools by experiencing empowerment, recognition, and an improved quality of life. The integration of multiple resources in one technological tool supports caregiving, saving time and facilitating daily tasks. The collaboration of ICs in technology development is crucial for innovation and overcoming barriers, emphasizing the need for user-driven solutions.

**Limitations**

As for the analysis of the included studies, it was not possible to use a tool to evaluate study quality. In the papers that were analyzed, it was not possible to identify references about improvement opportunities arising from the research process. The fact that the samples in the presented studies are small does not allow us to extrapolate data to the population.

The included papers were published in English, French, Spanish, and Portuguese, and the inclusion of articles in other languages could have brought more relevant information to this review. However, searching in 4 databases allowed us to expand the search comprehensiveness.

**Comparison With Previous Work**

In Portugal, there are few scientific studies carried out by nurses that refer to the use of digital technologies as a resource to train people with dependence and ICs.

**Conclusions**

Evidence found in studies revealed that ICTs such as digital platforms, portals, and web-based community groups were preferentially used by informal caregivers via mobile apps and that computers were used more by the people cared for. Studies showed that ICs had access to and used digital technology not only to meet the needs of the person cared for but also to meet their own needs. Studies have shown that digital technology is an accessible tool for empowering ICs. However, there were concerns regarding privacy, security, and the use of these tools, which should be considered by health care professionals and researchers. It is also important to highlight the necessity of providing digital training for both ICs and the individuals under their care.

ICs play a key role in the provision of quality care to the dependent people to whom they commit. It is crucial to understand how digital tools can be effectively and beneficially used to empower ICs.

The participation of ICs is essential when it comes to developing digital tools (platforms, mobile apps, and portals) because they can contribute to developing tools that meet users’ needs (ICs and the people cared for). The use of digital technologies can guarantee access to knowledge, thereby empowering caregivers when it comes to making a decision and sharing care provision with health care professionals. It is important to emphasize the significance of digital empowerment in enhancing the digital health literacy of both ICs and those they care for. Digital technology allows accessible, targeted, and effective communication. Health care professionals and researchers should guarantee information reliability, security, and clarity and optimize existing resources.

**Authors’ Contributions**

SS contributed to the design of the paper; contributed to the collection, analysis, and interpretation of the data; and drafted the paper. LVH contributed to the interpretation of the data, drafted the paper, and substantively revised it. MFF contributed to the design of the paper and the collection, analysis, and interpretation of the data. AH contributed to the design of the paper and the interpretation of the data. GC contributed to the interpretation of the data and substantively revised it. ASC contributed to the design of the paper and the interpretation of the data and drafted and substantively revised the manuscript.

**Conflicts of Interest**

None declared.

Multimedia Appendix 1
PRISMA-ScR checklist.

[PDF File (Adobe PDF File), 350 KB - aging_v7i1e54913_app1.pdf ]

**References**

2. Hassan AY. Challenges and recommendations for the deployment of information and communication technology solutions for informal caregivers: scoping review. JMIR Aging 2020 Jul 29;3(2):e20310 [FREE Full text] [doi: 10.2196/20310] [Medline: 32723720]


Abbreviations

HL: health literacy
IC: informal caregiver
ICT: Information and Communication Technology
JBI: Joanna Briggs Institute
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
WHO: World Health Organization
Original Paper

Digital Adoption by an Organization Supporting Informal Caregivers During COVID-19 Pandemic Showing Impact on Service Use, Organizational Performance, and Carers’ Well-Being: Retrospective Population-Based Database Study With Embedded User Survey

Ala Szczepura1*, BA, MA, DPhil; Amir Jahan Khan2*, BSc, MSc, PhD; Deidre Wild1*, BA, RGN, MSc, PhD; Sara Nelson3*, BSc, PhD; Sonja Woodhouse4*, CIPD, PGDip, MSc; Mark Collinson5, MA, MBA

1Research Centre for Healthcare & Communities, Coventry University, Coventry, United Kingdom
2Department of Economics, Institute of Business Administration (IBA), Karachi, Pakistan
3Prostate Cancer Research, London, United Kingdom
4Carers Trust UK, London, United Kingdom
5MC2S Consultancy Services, Bromsgrove, United Kingdom
*these authors contributed equally

Corresponding Author:
Ala Szczepura, BA, MA, DPhil
Research Centre for Healthcare & Communities
Coventry University
Richard Crossman Building
Jordan Well
Coventry, CV1 5RW
United Kingdom
Phone: 44 0 7557 425 463
Email: ab5794@coventry.ac.uk

Abstract

Background: The COVID-19 pandemic has catalyzed a move from face-to-face to digital delivery of services by hospitals and primary care. However, little is known about the impact of digital transformation on organizations supporting unpaid caregivers. Since the start of the COVID-19 pandemic, the value of care provided by such informal caregivers is estimated to be £111 billion (US$ 152.7 billion) in England.

Objective: This study aims to analyze service uptake patterns (including digital service options) over the pandemic period in an English caregivers’ support organization covering a population of 0.98 million; measure changes in organizational performance, service efficiency, and quality; and identify the views of caregivers on service provision and future digital delivery.

Methods: This was a retrospective analysis of the use of digital versus nondigital support services (January 2019 to June 2021) by caregivers in city and rural geographic areas. We compared organizational performance and service quality indicators for 2 financial years (2019-2020 and 2020-2021). A survey was conducted to identify barriers and facilitators to digital service uptake, the computer proficiency of caregivers (the Computer Proficiency Questionnaire, 12-item version), and preferences for future digital service provision. Quantitative data were analyzed using Stata 13 (StataCorp LLC). Thematic analysis was used for open-text survey responses.

Results: The number of caregivers registered with the organization rose from 14,817 in 2019 to 20,237 in 2021. Monthly contacts rose from 1929 to 6741, with remote contacts increasing from 48.89% (943/1929) to 86.68% (5843/6741); distinctive patterns were observed for city versus rural caregivers. There was an increase in one-to-one contacts (88.8%) and caregiver assessments (20.9%), with no expansion in staffing. Service quality indicators showed an improvement in 5 of 8 variables (all P<.05). The 152 carers completing the survey had similar demographics to all registered caregivers. The Computer Proficiency Questionnaire, 12-item version, mean score of 25.61 (SD 4.40) indicated relatively high computer proficiency. The analysis of open-text responses...
identified a preference for the organization to continue to offer face-to-face services as well as web-based options. The digital services that were the most highly rated were carers’ well-being assessments, support needs checks, and peer support groups.

Conclusions: Our findings show that staff in the caregiver support organization were agile in adapting their services to digital delivery while dealing with increased numbers of registered clients and higher monthly contacts, all without obvious detriment to service quality. Caregivers indicated a preference for blended services, even while recording high computer proficiency. Considering the economic importance of unpaid caregivers, more attention should be given to organizations funded to provide support for them and to the potential for technology to enhance caregivers’ access to, and engagement with, such services.

(JMIR Aging 2024;7:e46414) doi:10.2196/46414

KEYWORDS
digital adoption; COVID-19; informal caregivers; carer support organization; organization performance; integrated care systems; care systems; health policy; aging in place; digital divide

Introduction

Background

During the COVID-19 pandemic, social distancing policies limiting physical contact transformed how people were able to access health and care services and sped up digital transformation in many sectors. In the United Kingdom, traditional hospital services rapidly moved from face-to-face to digital services, from simple websites to web-based platforms, to reduce the chance of infection transmission [1]. In primary care, general practitioners (GPs) also adopted a digital first approach for consultations [2]. However, how those organizations that are contracted to support unpaid caregivers adapted their services is not known [3]. Such organizations are not part of the National Health Service (NHS) but are instead contracted by local government [4]. Before the pandemic, expenditure on the services these organizations provided was reduced by 11% over the 6-year period from 2015 to 2021 [5], while 36,000 more carers were directed to their services [5]. It is recognized that unpaid carers play a crucial role in providing essential care worldwide [6]. Bearing these facts in mind, the absence of research on how these organizations moved to web-based delivery, what impact this had on service delivery, their ability to do this without affecting service quality, and the response of client caregivers are noteworthy.

Before the pandemic, an NHS report on the Widening Digital Participation Programme identified that organizations supporting the well-being of carers had largely been forgotten [7]. Currently, many countries are introducing policies of active aging, supporting older people to live independently in the community rather than entering expensive long-term residential care [8,9]. As a result, while the number of staff employed by care homes has remained unchanged from 2012 to 2022 in England, there has been a 27% increase in domiciliary care workers who support people in their own homes [10]. The pressures on unpaid carers have inevitably increased. Currently, the care provided by unpaid carers in England is estimated to be equivalent to that provided by 3.2 million full-time paid care workers [11,12], nearly 6 times the size of the paid workforce of 510,000 domiciliary care workers [13]. In addition, the continuity of care provided by family carers is becoming increasingly important in the context of a high annual turnover of domiciliary care staff (31.5% in England) [13]. Across the world, countries are developing integrated long-term care strategies to support their aging populations as recommended by the World Health Organization [14]. England has established 42 new integrated care systems (ICSs) to underpin integration between health and care services in these geographic areas [15]. Proposals for joining up care include the idea of wrap-around services for care recipients plus their caregivers, but there is no mention of organizations that provide support for unpaid caregivers. Although the national ICS strategy incorporated a digital plan, this currently excludes mention of organizations that support informal carers [16]. Such services are typically provided by charities or not-for-profit organizations that themselves may have limited expertise in digital transformation [17]. In this context, it is important to better understand the experiences of such carer support organizations during the pandemic and the response of the caregivers they support.

Objectives

This study aims to provide evidence to address this important research gap. The research has 3 main objectives: to analyze changes in service use patterns (including services accessed and the uptake of digital options); to assess any impact on organizational performance, service efficiency, and key quality indicators; and to identify user clients’ preference for future digital services.

Methods

Overview

We analyzed data collected by an organization providing support for 20,237 caregivers, covering city and rural geographic areas. Data were downloaded and fully anonymized. The uptake of digital and traditional services was examined over a 30-month period from January 2019 (before the pandemic) to June 2021 (after COVID-19-related restrictions were lifted). Service-level performance and proxy quality indicators were constructed and compared for the prepandemic financial year (2019-2020) and through the following initial lockdowns (2020-2021). The analysis of a user feedback survey undertaken at the end of this period (September 2021) explored barriers and facilitators to digital service uptake, the computer proficiency of caregivers, and views on future digital service provision. The organization’s digital preparedness before the pandemic was assessed.
Ethical Considerations
This retrospective study using fully anonymized existing data received ethics approval from the Coventry University Ethics service (P163079).

Organizational Setting
The study was undertaken at the Carers Trust Heart of England (CTHE). The organization operates in the complex UK sector that provides support for unpaid carers [4]. The CTHE is contracted to provide carer assessments and caregiver well-being services in 2 separate geographic areas: city area (Coventry) and rural area (Warwickshire, including towns and villages). The 2 areas have a total population of 0.98 million, and they are covered by a single ICS. Services provided include an assessment of a carer’s needs, information on health and care services, benefit entitlement, assistive technologies, and peer support. Before the pandemic, the CTHE had made a number of changes to its IT systems, moving everything to a single cloud platform so that databases could be accessed from anywhere. Laptop computers and mobile phones had been provided to all frontline staff, all of which greatly facilitated home working. Job descriptions were also adapted to mention hybrid working. Throughout the observation period, the number of CTHE staff providing direct support to adult carers remained relatively constant at 14 well-being advisers, 3 administrators, 3 specialist roles (ethnic minority support worker, mental health worker, and carer trainer), and a manager for each area. A separate team provided support for young carers; this activity was excluded from our study. The CTHE is a member of a network of 124 Carers Trust partners across England, Scotland, and Wales. Members work within a national framework of policies, procedures, and internal quality assurance programs.

Longitudinal Data and Analysis
Overview
CTHE staff downloaded select activity data routinely collected for adult carers registered with the service (Multimedia Appendix 1). Young carers (aged <18 y) were excluded. Data were fully anonymized before being provided to the research team for analysis. Data cleaning and analysis were conducted on the imported raw data using reproducible coding files. Statistical analysis used Stata 13 (StataCorp LLC). The hypothesis was that there would be an increase in the use of web-based methods, although it was unclear whether digital levels would be sustained, what the impact on service delivery levels and quality would be, or what views client carers might have on a future digital service.

Service Use Patterns
Descriptive statistics were used to examine service use patterns and changes in the means of accessing services (ie, in person vs digital) [18]. Monthly contacts handled by the service were the primary variable used to explore use patterns over the 30-month period (January 2019 to June 2021). This covered two 15-month pre- and postpandemic periods, with the first national lockdown occurring midway in March 2020. The levels of service use by city versus rural carers were examined, together with the primary reasons for contacting the CTHE.

Service Performance Levels and Quality Indicators
Four key activity measures reported to commissioners each financial year were extracted from the data downloads. These included the number of carers supported, the number of one-to-one contacts, the number of carers’ assessments completed, and numbers of carers attending group activities. The CTHE also collected structured feedback from the client after every contact (Multimedia Appendix 1). Five proxy quality indicators routinely reported to the funder were also extracted. A further 3 proxy quality indicators were constructed from the raw data to identify whether a contact had reduced stress, increased control of personal life, or increased confidence. Changes in all 8 quality indicators were compared for the 2 financial years 2019-2020 and 2020-2021, with P values for percentage changes estimated using the Pearson chi-square test of association between the characteristic variables and the corresponding totals [19].

Survey of Client Caregivers
Registered adult carers were invited by the CTHE to complete a feedback questionnaire in September 2021 to review provision and help plan future services (Multimedia Appendix 2). Responses were fully anonymized before being provided to the research team, with a minimum target of 60 responses set for analysis. Closed questions requested information on the use of digital services, barriers and facilitators to access, and satisfaction with the services accessed, while open-text comment boxes enabled people to expand on their responses. Respondents were also invited to complete the Computer Proficiency Questionnaire, 12-item version (CPQ-12) to assess their computer proficiency [20]. As well as being made available on the web, feedback questionnaires were also mailed by the CTHE to carers. The characteristics of respondents were compared to those of adult carers on the register where possible. Open-text responses were analyzed for content using thematic analysis and cross-referenced to gain an understanding of the underlying reasoning behind the views expressed [21].

Results
Longitudinal Data Analysis
Registered Carer Characteristics
Data on registered carers (Multimedia Appendix 1) were analyzed over the 30-month period from January 2019 to July 2021. During this time, the total number of carers registered with the organization increased from 14,817 to 20,237 (Table 1). Those living in the rural area rose from 14,817 to 20,237 (18.3% increase), while those in the city showed a more modest rise from 13,132 to 15,459 (17.7% increase). Overall, the proportion of rural caregivers increased from 1 in 10 (1685/14,817, 11.37%) to 1 in 4 (4778/20,237, 23.61%).
Table 1. Demographics of registered caregivers and survey respondents.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Caregivers, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Rural caregivers</strong></td>
<td></td>
</tr>
<tr>
<td>January 2019^a</td>
<td>1685 (11.37)</td>
</tr>
<tr>
<td>January 2020^b</td>
<td>3035 (17.63)</td>
</tr>
<tr>
<td>July 2021^c</td>
<td>4778 (23.61)</td>
</tr>
<tr>
<td><strong>Registered caregivers (n=17,641; March 2020)^d</strong></td>
<td></td>
</tr>
<tr>
<td>Race and ethnicity</td>
<td></td>
</tr>
<tr>
<td>Black (Caribbean or African)</td>
<td>406 (2.3)</td>
</tr>
<tr>
<td>South Asian</td>
<td>2382 (13.5)</td>
</tr>
<tr>
<td>White</td>
<td>13,972 (79.2)</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>11,467 (65)</td>
</tr>
<tr>
<td>Aged ≥65 years</td>
<td>9067 (51.4)</td>
</tr>
<tr>
<td><strong>Employment status</strong></td>
<td></td>
</tr>
<tr>
<td>Retired or gave up work to care</td>
<td>7868 (44.59)</td>
</tr>
<tr>
<td>Working or in training</td>
<td>3369 (19.1)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>1270 (7.2)</td>
</tr>
<tr>
<td><strong>Survey respondents (n=152; September 2021)</strong></td>
<td></td>
</tr>
<tr>
<td>Rural caregivers^e</td>
<td>80 (56.3)</td>
</tr>
<tr>
<td>Sex (female)</td>
<td>97 (63.8)</td>
</tr>
<tr>
<td><strong>Employment status</strong></td>
<td></td>
</tr>
<tr>
<td>Retired or gave up work to care^f</td>
<td>33 (59)</td>
</tr>
<tr>
<td>Working or in training^f</td>
<td>18 (32)</td>
</tr>
<tr>
<td>Unemployed^f</td>
<td>2 (3)</td>
</tr>
</tbody>
</table>

^aTotal registered carers: 14,817.
^bTotal registered carers: 17,246.
^cTotal registered carers: 20,237.
^dMidway through the study period.
^e142 respondents provided information on their location.
^f56 respondents provided information on their employment status.

The demographic characteristics of registered carers were analyzed based on data recorded midway through the observation period from January 2019 to March 2020 (Table 1). The age and sex breakdown is comparable to national figures for carers [12]. Two-thirds (11,467/17,641, 65%) of the respondents were female carers; approximately half (9067/17,641, 51.4%) were aged ≥65 years; and over three-quarters (13,972/17,641, 79.2%) were White. In terms of their employment status, nearly half (7868/17,641, 44.59%) were retired or had given up work to care, with only 1 in every 5 (3369/17,641, 19.1%) working or in training.

**Service Use Patterns**

Over the 30-month period (January 2019 to July 2021), digital services were provided alongside face-to-face services. The former included the use of email, SMS text messaging, Zoom, WhatsApp, and Microsoft Teams, as well as web-based group sessions. Use patterns were analyzed over time (Figure 1). The 3 vertical lines indicate the time points at which major national COVID-19–related restrictions were applied (ie, lockdown periods). From 2019 to 2021, monthly carer contacts with the well-being support service rose from 1929 to 6741, with telephone contacts rising from 818 to 3071 per month, and digital contacts from 125 to 2772 per month. A separate analysis of digital contacts during this period (Figure 1A) shows that, alongside a near–5-fold overall increase in the monthly rate, there were peaks coinciding with the national lockdowns. A separate analysis of digital versus nondigital contacts (Figure 1B) uncovers a clear change in the balance between the two, with digital contacts rising from 6.48% (125/1929) to 41.12% (2772/6741) of all contacts by the end of the period.
A separate examination of rural and city caregivers uncovers distinctly different patterns. For city carers (Figure 1C), digital contacts started to climb steadily from the start of the pandemic, almost reaching parity with face-to-face contacts around the time of the second lockdown. They then started to tail off toward the end of the 30-month observation period. Rural carers (Figure 1D) demonstrated a much slower initial uptake, with the rate of digital adoption only really starting to pick up after the second and third lockdowns. However, unlike in the case of city carers, rates were continuing to rise at the end of the observation period.

**Types of Digital Contacts Used**

Digital contacts were categorized into 4 broad groups. Two represented more flexible asynchronous methods (ie, email and SMS text messaging), whereas 2 represented fixed-time synchronous methods (ie, internet-based communication using Zoom, WhatsApp, and Microsoft Teams) and social digital group activities. Before the pandemic, only asynchronous methods were used, and rates were very low (Figure 2). With the first national lockdown (month 15), the use of other digital methods started to be added. Over time, the use of internet-based communication increased, overtaking SMS text messaging, although email remained the principal form of contact. “Social” groups were the least used form of digital contact used, possibly due to greater difficulties in arranging and delivering these. Their use peaked between September 2020 and January 2021 (ie, between the second and third lockdowns).
Types of Services Accessed

Well-being services available to carers at the start of the pandemic were mostly delivered via carers center appointments, home visits, and various outreach activities. During the pandemic, carers center appointments ceased after the first lockdown, with home visits and outreach activities significantly reduced. The well-being services provided were information and advice on aids and adaptations, carers’ assessment, education and training (including digital), emotional support, local and national health and social care services, personal care and health, residential care and day care (for cared-for persons), and social inclusion and interests (including links to local and national groups). Other services included peer support and emergency planning. During the pandemic, additional services were introduced to address specific needs: a grief and loss service, employment support working collaboratively with employers, and finance support to reduce carer hardship.

Analysis of the reasons for contacting the well-being service over the 3-year period from 2019 to 2021 indicated that, irrespective of the year, the 3 most common reasons were emotional support, local and national information, and personal care and health (Figure 3). All 3 reasons peaked in 2020 (early pandemic period), as did information on day care relief and aids and adaptations. Advice related to digital inclusion and social inclusion were unusual in showing a steep rise in the final 6 months, from a base where both were 0 in 2019. Education and training showed a similar pattern, possibly partly related to digital training needs. The most frequently new service accessed was finance, with the grief and loss service reaching a similar level in the final 6 months.

Figure 2. Number of digital contacts per month versus type of remote access.
Service Performance and Quality

Activity levels demonstrated an increase of 88.8% over 2 years in the numbers of one-to-one contacts and an increase of 20.9% in individual carer assessments (Table 2). At the same time, there was a decrease of 70.6% in numbers attending group activities. Among the 5 proxy quality indicators routinely reported to the funder, those critical to well-being all showed an improvement (ie, feeling less alone, making it easier to cope with caring role, and helping improve physical health), as did the extra 3 indicators constructed (ie, reduced stress, increased control of personal life, and increased confidence). Improvements ranged from 13% for reduced stress ($P=0.001$) to 1% for others ($P>0.05$). The only measure to show a significant decrease was help dealing with health and social care professionals ($P=0.001$). This may, in part, be due to pressures experienced by health and social care staff during the pandemic.
### Table 2. Change in carer contact type and quality indicators (from 2019-2020 to 2020-2021).

<table>
<thead>
<tr>
<th>Carer contact type</th>
<th>2019-2020</th>
<th>2020-2021</th>
<th>Change (%)</th>
<th>P value(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-to-one contacts with carers, n</td>
<td>7344</td>
<td>13,868</td>
<td>+88.8</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Carers supported on one-to-one basis, n</td>
<td>3021</td>
<td>4463</td>
<td>+47.7</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Carer assessments completed, n</td>
<td>234</td>
<td>283</td>
<td>+20.9</td>
<td>.03</td>
</tr>
<tr>
<td>Carers attending group activities, n</td>
<td>670</td>
<td>197</td>
<td>−70.6</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Registered caregivers, n</td>
<td>14,817(^b)</td>
<td>17,641(^c)</td>
<td>+19.1</td>
<td>N/A(^d)</td>
</tr>
</tbody>
</table>

**Benefits reported after contact**

**Helped me feel less alone in my caring role**
- Contacts, n: 2249 to 5217, change +131.0, P <.001
- Contacts (%): 33 to 36, change +9.0, N/A
- Caregivers expressing this, n: 1134 to 2410, change +112.5, P <.001

**Helped me reduce my stress**
- Contacts, n: 1771 to 4378, change +147.2, P <.001
- Contacts (%): 18 to 31, change +72.2, N/A
- Caregivers expressing this, n: 948 to 2410, change +154.2, P <.001

**Made it easier to cope with my caring role**
- Contacts, n: 825 to 1319, change +59.9, P <.001
- Contacts (%): 8 to 9, change +12.5, N/A
- Caregivers expressing this, n: 597 to 879, change +47.2, P = .95

**Helped me improve my physical health**
- Contacts, n: 465 to 1278, change +174.8, P <.001
- Contacts (%): 5 to 9, change +80.0, N/A
- Caregivers expressing this, n: 293 to 660, change +125.3, P <.001

**Helped me improve my financial position**
- Contacts, n: 692 to 804, change +16.2, P <.001
- Contacts (%): 7 to 6, change −14.3, N/A
- Caregivers expressing this, n: 469 to 513, change +9.4, P <.001

**Helped me deal with health and social care professionals**
- Contacts, n: 342 to 250, change −26.9, P <.001
- Contacts (%): 3 to 2, change −33.3, N/A
- Caregivers expressing this, n: 230 to 174, change −24.3, P <.001

**Helped me increase control of my personal life**
- Contacts, n: 305 to 638, change +109.2, P = .15
- Contacts (%): 3 to 4, change +33.3, N/A
- Caregivers expressing this, n: 218 to 350, change +60.6, P = .35

**Helped me increase my confidence**
- Contacts, n: 205 to 417, change +103.4, P = .39
- Contacts (%): 2 to 3, change +50.0, N/A
- Caregivers expressing this, n: 144 to 272, change +89.9, P = .02

\(^a\)Pearson chi-square test.
\(^b\)January 2019.
\(^c\)March 2020.
Survey of Client Caregivers

Respondents

A total of 152 caregivers completed the survey. Personal characteristics (age and sex) were largely comparable to those of all caregivers registered with the CTHE (Table 1). Ethnicity was not recorded in the survey. In terms of employment, the 2 largest groups were those who had retired or people who had given up work to care and those in work or training. The survey respondents included a higher percentage of rural caregivers than the carer register (85/142, 59.9% vs 4778/20,237, 23.61%); of the 152 respondents, 10 (6.6%) did not provide information on their location. Information provided in the survey showed that nearly half of these respondents (70/152, 46.1%) were caring for a husband, wife, or partner and 27.6% (42/152) for a parent, whereas the remaining 26.3% (40/152) had another relationship. Nearly one-third (41/148, 27.7%) of people providing more detailed information were caring for someone who was living on their own, and in 78.9% (112/142) of responses, care was provided solely by the unpaid caregiver. Where there was access to support from domiciliary carers, this averaged 19.5 (SD 30.7; range 2-40) hours per week. The CPQ-12 Questionnaire was completed by 94 (61.8%) of the 152 respondents. Respondents achieved a mean score of 25.61 (SD 4.40), with city and rural carers exhibiting similar proficiencies (Multimedia Appendix 3 [20,22]).

Prepandemic Use of Services and Main Barriers

Historically, respondents reported using a broad range of services (Table 3). Reasons for not accessing a particular service were provided by 82.2% (125/152) of the respondents; the remainder (27/152, 17.8%) reported only registering with the CTHE after the first lockdown. Two barriers to use were timing and travel distances, but far more respondents, from 28.6% (30/105) to 39.6% (40/101), perceived lack of awareness of a service as a barrier to its use. A further 33% (30/90) to 44.6% (45/101) stated that, before the pandemic, they had no need for a particular service. During the pandemic, some respondents had stopped using community outreach sites (14/114, 12.3%) and carers centers (19/114, 16.7%); a very small percentage (2/114, 1.8% to 4/114, 3.5%) had started to use these services. With cessation of home visits, 21.9% (25/114) reported that they had started to use telephone support during the pandemic.

<table>
<thead>
<tr>
<th>Service</th>
<th>Used 1-12 times (2019), n/N (%)</th>
<th>Barriers to use of a service, n/N (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Timing not suitable</td>
<td>Too far to travel</td>
</tr>
<tr>
<td>Carers center (drop-in visit)</td>
<td>23/116 (19.8)</td>
<td>12/105 (11.4)</td>
</tr>
<tr>
<td>Carers center (appointment visit)</td>
<td>16/117 (13.7)</td>
<td>8/101 (7.9)</td>
</tr>
<tr>
<td>Home visits</td>
<td>16/112 (14.3)</td>
<td>9/101 (8.9)</td>
</tr>
<tr>
<td>Outreach (eg, community site)</td>
<td>14/114 (12.3)</td>
<td>4/101 (4)</td>
</tr>
<tr>
<td>Telephone support</td>
<td>48/122 (39.3)</td>
<td>4/88 (4.5)</td>
</tr>
<tr>
<td>Group activities</td>
<td>34/121 (28.1)</td>
<td>11/90 (12.2)</td>
</tr>
</tbody>
</table>

Which Digital Services Were Used During the Pandemic?

Three-quarters (114/152, 75%) of the respondents identified (from a predefined list) which digital support services they used during the pandemic. Email (58/114, 50.9%), Zoom or Microsoft Teams (32/114, 28.1%), and WhatsApp, SMS text messaging, or video (10/114, 8.8%) were most commonly used, mirroring data presented in Figure 2. No respondent had used Skype or FaceTime. The most valued digital support services were carers’ well-being assessments, support needs checks, and peer support groups. When asked to indicate which web-based group activities they had experienced (respondents could tick as many as appropriate), 28.8% (34/118) replied. Those carers who provided a response most frequently accessed web-based training and resilience courses (14/24, 58%), virtual yoga sessions or quizzes (10/23, 43%), virtual cafes (9/22, 40%), and “carers evening chat” (4/16, 25%).

A total of 102 comments were entered by participants. These were analyzed for thematic content and commonalities. Three superordinate themes were identified (Textbox 1). These included 8 subthemes (2-3 subthemes emerged for each superordinate theme). A selection of comments relating to each subtheme were extracted. The first theme, how to help carers use digital services in the future, highlighted aspects such as a need for more publicity, activities provided at different times and in different formats, and technical help for persons who are digitally excluded. The second theme, offering a selection of well-being services, contained 2 strong subthemes. One was the view that digital services are invaluable and the second was that face-to-face services are essential for certain functions and for those who are digitally excluded because of their age or for financial reasons. The third and final theme emerging from users was the need to tailor future digital services to meet individual caregivers’ needs. This might include, for example, not only addressing the practical elements of caring or issues of isolation and confidentiality but also acknowledging that “people need more than their problems fixing” and not losing the personal service previously provided, which is highly valued.
### Theme 1: things to help me use digital services in future

- The need for more publicity and better communications
- "I was not aware of the online services, more information/publicity would be helpful."
- "Better information of services you can use."
- "Old fashioned ‘come join us’ flyer through the post. My elderly parents need constant encouragement and they don’t read emails and my repeating them...or reading them out does not have the same impact."
- Technical help and digital inclusion
- “[C]an’t use a computer, [need] help to set up Zoom.”
- “I am not comfortable using online services.”
- “Not everyone is online so cannot avail themselves to online forums...also the cost of broadband needed to use Zoom etc. which requires higher speed etc. is quite prohibitive if you are living on a fixed income.”
- Activities at different times and in different formats
- “They just don’t fit in as they are at times when I’m caring for my mother.”
- “Due to work pattern [I am] not always available.”
- “A wider range of subjects for the online service. Perhaps short podcasts of interesting places in the world e.g. videos of museums around the world or tourist destinations or cultures and traditions or other countries.”
- “A blend of online and face-to-face better. Also, a brief catch-up call if you can’t make a session as guilt can set in for me if I’m overloaded and I feel unable to continue.”

### Theme 2: offering a selection of well-being services, including on the web

- Web-based services are invaluable
- “Online saves time travelling and you can access it whilst still caring for the patient in your own home.”
- “Living through COVID has been like being relocated to the moon, no contact with anyone. At a time when in person is still beset with logistical problems the online equivalent is a lifeline.”
- Face-to-face services still necessary
- “It (digital) was a necessary substitute during lockdown but nothing replaces face-to-face interaction.”
- “Nothing is as helpful as face-to-face help, especially where counselling and support services are concerned. Many carers, especially those caring for someone with dementia are elderly and not used to computers.”

### Theme 3: tailoring future services

- Practical elements of caring
- “A list of possible areas to look at and their contact details e.g. home cleaners, meals on wheels etc.”
- “I have as a parent, many worries about what will happen when I can’t ‘go on’...Legal advice for preparations for the outcome.”
- “Recommended places approved by members experiences. Where to get...Equipment, grants assistance etc. positive recommendations for work carried out for adaptations by local companies.”
- Addressing isolation and confidentiality
- “I feel very isolated as a carer and being able to go to meetings/events/social gathering and see and speak to people normally I feel is very important, both for me as a carer, and my husband who has mixed dementia.”
- “I feel it is of great importance that you can discuss on a one-to-one level in person or telephone on the day. Not every carer can talk freely about what is going on for themselves and especially if the cared for is listening.”
- Addressing changing times and loss of personal service
- “Go back 15 years and the carers center was a place where you could turn up to have a chat with whoever was on the desk. With the move to the library the feel changed—interactions more like ‘please state the nature of your problem’ than ‘how are you, how are things going?’ People need more than their problems fixing...it is more the emotional and community support. That’s it—emotional support as well as practical support.”
Discussion

Principal Findings

To our knowledge, this is the first large-scale study to analyze the impact of the COVID-19 pandemic on the provision of well-being services for unpaid caregivers. The longitudinal analysis of >20,000 rural and city carers identified, as expected, a move away from face-to-face to web-based service access. The shift observed mirrors those reported by researchers for other services during the pandemic, such as GP practices [2]. However, over the period from January 2019 to June 2021, the number of monthly contacts handled by the carer support organization more than tripled, with no significant changes in staffing levels. Before the pandemic, digital contacts were exclusively by email or SMS text messaging. During the pandemic, additional options were introduced, including Zoom, WhatsApp, and Microsoft Teams, with new web-based group sessions also offered to carers.

Within the context of a 37% increase in the number of carers registered, the organization managed to increase the number of one-to-one contacts by 88.8%. This increase was particularly evident in rural areas, with the ratio of such carers on the register rising from 1 in 10 carers to 1 in 4 carers by the end of the observation period. Rural carers demonstrated slightly different behaviors, showing much slower initial digital adoption rates. In terms of carers’ rating of the service received, 6 of the 8 quality indicators showed an improvement, and the other two showed only a minor, nonsignificant decrease. The largest improvement was observed in reduced stress, consistent with the findings of a systematic review of caregiver web-based interventions [23]. Our survey identified a high level of computer proficiency among carers, at or above that reported for other older populations [20,22]. Even so, respondents expressed a preference for the organization to continue to offer face-to-face services as well as web-based options to meet a carer’s preferences and the type of well-being support required.

Comparison With Prior Work

Before the pandemic, researchers reported that the uptake of web-based services by older adults in the United Kingdom remained relatively low, despite their potential benefits [24]. A European examination of web-based services available to support informal carers also found a lack of reliability and usability [25]. A qualitative study of the views of caregivers on suitable technologies to assist their caregiving identified similar themes to this larger study, in particular that digital technology needs to be tailored to users’ needs in order to ensure adoption [26]. Although it is acknowledged that there may be a huge potential to use such tools to support unpaid carers, it is recognized that wholesale adoption may risk inadvertently exacerbating existing support through digital exclusion [27]; for example, the testing of digital tools in a real-world setting has identified a digital inverse care law, with those most in need of support least likely to engage with digital health platforms [28]. In addition, a review of eHealth interventions to support caregivers of older adults also highlights the importance of using appropriate language and text, as well as helping caregivers learn how to use the intervention [29].

Our research shows that, during a crisis such as the COVID-19 pandemic, an organization providing support for the well-being of caregivers was able to successfully implement remote service provision using a mix of traditional and digital tools without a detrimental impact on the reported quality of individual contacts and in the context of an increased workload. Systematic reviews of internet-based interventions to support caregivers have to date reported mixed results and called for more high-quality studies [30,31]. A recent review of factors influencing the implementation of eHealth to support informal care found a gap in knowledge regarding success factors and limited focus on the well-being of the unpaid carer, with the focus being principally on the person receiving care [32]. Similarly, studies of telecare that focus on conditions such as dementia usually do not differentiate the caregivers’ needs, instead usually considering the caregiver and the older person or care recipient as a dyad [33-35]. Early in the pandemic, there were some calls to move “carers from the back of the queue” when considering digital services [36]. However, a recent research study of digital interventions for carers of people with dementia still considers need in terms of the dyad, with caregivers in a secondary role [37]. In the United Kingdom, the 2019 report for government on preparing the health care workforce for the digital future recommended that the NHS should work with carer organizations to prioritize the education of patients and caregivers alongside the health care workforce [38].

It is important to consider the indicator that showed a significant deterioration in our study. This was associated with support in accessing health care and social care services, both presumably disrupted by the COVID-19 pandemic. A review of carer support has identified that the ability to coordinate access to such services is particularly valuable, with the integration of home care and community care able to improve outcomes for older people [39,40]. This seems to be particularly important for carers of people with a mental health condition [41]. In Australia, the integration of digital care and clinical care is being assessed to coordinate mental health teams, caregivers, and service users as active partners [42]. The potential for appropriate digital technology to provide support and reassurance is recognized as a benefit for both the caregiver and the person for whom they care [43]. In some parts of the world, volunteers are also being integrated into care to help caregivers use custom-built apps [44].

Investment in innovation to provide optimum support services for informal caregivers could be highly cost-effective. The workforce of unpaid carers represents approximately 6% of the UK population and, together with the 1.3 million registered carers who receive a small carer’s allowance, informal caregivers are widely acknowledged as a crucial component in care delivery [12,45]. Furthermore, since the start of the pandemic, the value of unpaid care provided in England is estimated to be £111 billion (US $152.7 billion) [46] and in the United States >US $450 billion annually [47]. The UK government has recently set out a range of policies aimed at empowering unpaid carers, with a dedicated, although small, £25 million (US $26.72 million) budget for this purpose [46]. The danger is that the initiatives will once again focus on caregivers rather the sector that supports them. Thus, the...
opportunity to integrate organizations such as the one in this study into the wider community-based care system will be missed. This study also questions the stereotype of low digital capability among older carers, with CPQ-12 scores demonstrating high computer proficiency. Even so, survey responses indicate that large-scale naive digital transformation is unlikely to be effective. Instead of a “one-size-fits-all” approach, there is a need for person-centered support (face-to-face as well as web-based options) as part of the service, together with training for those who need it. Meanwhile, there are emerging indications of a move toward providing solely web-based support, with some suppliers looking ahead to younger and more digitally engaged carers who are assumed not to require face-to-face services [48]. The lower cost of web-based support services may make this seem an attractive option for commissioners in the United Kingdom. A similar situation has occurred in primary care with disruptive innovators entering the NHS market to provide web-based GP services, with the evaluation reporting mixed findings and providers withdrawing from some NHS contracts [49]. For any caregiver support service evaluation, as well as delivery costs, there will need to be a careful consideration of utility (ie, quality and effectiveness from the user perspective) [50]. Although this study was set in England, the findings will be relevant for other countries where digital services to support the well-being of informal caregivers are in use or are being developed.

Limitations
There are a number of limitations to this study that need to be acknowledged. First, it is unclear how representative, in terms of its digital readiness, the organization studied is of the whole sector. There are no national audits of such carer support organizations, although it is known that >60% of care homes still use internet connections that will not support full digital transformation [16]. Second, the cohort excluded young carers, which inevitably limits generalizability to the wider population of carers [51]. In addition, an important subgroup (working caregivers) could not be identified due to limitations in the data. A quarter of older workers in England currently have caring responsibilities, and this percentage is expected to increase as the population ages [52]. Third and last, although we identified a high level of computer proficiency in survey respondents, this may not be fully representative because most of the respondents (136/152, 89.5%) completed the survey on the web. Further research is needed to provide evidence on these subgroups before drawing any final conclusions about web-based support services.

Conclusions
Looking to the future, the integration of health care and care services to meet the complex care needs of a country’s aging population is recognized as a global challenge [53]. Considering the importance of unpaid carers, more attention needs to be given in all national strategies to organizations that support this important free workforce. Our study highlights a number of issues worthy of further consideration and study that have implications for the design of future cost-effective digital initiatives. These include the lack of any audits of the digital readiness of organizations that provide support for caregivers; the need for a better understanding of rural carers; evidence of the cost-effectiveness as well as the use of different forms of support for caregivers; and the potential for collaboration among different partners within ICSs to better support unpaid caregivers, enhancing their access to, and engagement with, support services after the COVID-19 pandemic.

Acknowledgments
The authors would like to thank a number of staff at Carers Trust Heart of England: Anna-Marie Shanley (quality and performance officer) for data downloads; Steve Thomson (head of business development), Yvonne Moore (business support service manager), and Jacqueline Lucid and Michelle Sawyer (well-being service managers) for valuable advice; and to Claire Dale, chief executive officer, and Melanie Boag, deputy chief executive officer, for their support. Thanks are due to Carers Trust Heart of England for funding the time of AK, SN, and MC to undertake the analyses. The funder had no role in study analysis, the interpretation of data and the writing of the manuscript, or in the decision to submit the paper for publication. Thanks are also due to Dr Michael Loizou (Wallscourt Associate Professor in Health Technology and Life Sciences, University of the West of England, Bristol, United Kingdom) who acted as adviser to the project.

Authors’ Contributions
MC and SW developed the original concept. MC, DW, and AS jointly developed the analysis plan for the study. AJK and SN undertook data analyses. AS wrote the first draft of the article, and all authors critically revised the paper for important aspects. All authors read and approved the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Metrics and outcomes recorded for adult carers registered with the well-being service.

[DOCX File , 21 KB - aging_v7i1e46414_app1.docx ]
References


6. Cotis JP. Population ageing: facing the challenge: ageing will make it hard for governments to deal with mounting financial pressures. It may be time to rethink our policies towards work. OECD Publications and Information Centre. 2003 Sep. URL: https://go.gale.com/ps/i.do?id=GALE%7CA110739426&sid=googleScholar&v=2.1&it=r&linkaccess=abs&issn=00297054&p=AONE&sw=w&userGroupName=anon%7Ee3388c9&atv=open-web-entry [accessed 2020-10-15]


Abbreviations

CPQ-12: Computer Proficiency Questionnaire, 12-item version
CTHE: Carers Trust Heart of England
GP: general practitioner
ICS: integrated care system
NHS: National Health Service

Edited by Y Jiang; submitted 10.02.23; peer-reviewed by M Wasilewski, W Tang; comments to author 18.04.23; revised version received 06.08.23; accepted 26.11.23; published 13.05.24.

Please cite as:
Digital Adoption by an Organization Supporting Informal Caregivers During COVID-19 Pandemic Showing Impact on Service Use, Organizational Performance, and Carers' Well-Being: Retrospective Population-Based Database Study With Embedded User Survey
JMIR Aging 2024;7:e46414
URL: https://aging.jmir.org/2024/1/e46414
doi:10.2196/46414
PMID:38739915
Sleep Duration and Functional Disability Among Chinese Older Adults: Cross-Sectional Study

Minjing Luo¹,²,³, PhD; Yue Dong⁴, MD; Bingbing Fan⁴, MD; Xinyue Zhang³, MD; Hao Liu¹,²,³, PhD; Changhao Liang¹,²,³, PhD; Hongguo Rong¹,²,³,⁵, PhD; Yutong Fei¹,²,³,⁵, PhD

¹Center for Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China
²Beijing GRADE Centre, Beijing, China
³School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China
⁴Dongzhimen Hospital Affiliated to Beijing University of Chinese Medicine, Beijing, China
⁵Institute for Excellence in Evidence-Based Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China

*these authors contributed equally

Corresponding Author:
Yutong Fei, PhD
Center for Evidence-Based Chinese Medicine
Beijing University of Chinese Medicine
No.11 Bei San Huan Dong Road, Chaoyang District
Beijing, 100026
China
Phone: 86 86 10 6428 6757
Email: feiyt@bucm.edu.cn

Abstract

Background: The duration of sleep plays a crucial role in the development of physiological functions that impact health. However, little is known about the associations between sleep duration and functional disability among older adults in China.

Objective: This study aimed to explore the associations between sleep duration and functional disabilities in the older population (aged ≥65 years) in China.

Methods: The data for this cross-sectional study were gathered from respondents 65 years and older who participated in the 2018 survey of the China Health and Retirement Longitudinal Study, an ongoing nationwide longitudinal investigation of Chinese adults. The duration of sleep per night was obtained through face-to-face interviews. Functional disability was assessed according to activities of daily living (ADL) and instrumental activities of daily living (IADL) scales. The association between sleep duration and functional disability was assessed by multivariable generalized linear models. A restricted cubic-spline model was used to explore the dose-response relationship between sleep duration and functional disability.

Results: In total, 5519 participants (n=2471, 44.77% men) were included in this study with a mean age of 73.67 years, including 2800 (50.73%) respondents with a functional disability, 1978 (35.83%) with ADL disability, and 2299 (41.66%) with IADL disability. After adjusting for potential confounders, the older adults reporting shorter (≤4, 5, or 6 hours) or longer (8, 9, or ≥10 hours) sleep durations per night exhibited a notably increased risk of functional disability compared to that of respondents who reported having 7 hours of sleep per night (all P<.05), which revealed a U-shaped association between sleep duration and dysfunction. When the sleep duration fell below 7 hours, increased sleep duration was associated with a significantly lower risk of functional disability (odds ratio [OR] 0.85, 95% CI 0.79-0.91; P<.001). When the sleep duration exceeded 7 hours, the risk of functional disability associated with a prolonged sleep duration increased (OR 1.16, 95% CI 1.05-1.29; P<.001).

Conclusions: Sleep durations shorter and longer than 7 hours were associated with a higher risk of functional disability among Chinese adults 65 years and older. Future studies are needed to explore intervention strategies for improving sleep duration with a particular focus on functional disability.

(JMIR Aging 2024;7:e53548) doi:10.2196/53548

KEYWORDS

sleep duration; functional disability; activity of daily living disability; instrumental activity of daily living; older population
Introduction

The issue of disability in the older population has garnered significant attention and interest in recent decades [1-3]. Disability—as defined by the International Classification of Functioning, Disability, and Health—encompasses the combined effects of impairments, activity limitations, or participation restrictions [4]. Functional disability is a significant measure of restrictions in activities, specifically referring to challenges in performing basic activities of daily living (ADLs) and/or instrumental activities of daily living (IADLs) [5,6]. ADLs are widely recognized as essential tasks for maintaining independence in one’s own residence, such as dressing, bathing, and eating. IADLs refer to more complex tasks that require a higher level of independence and cognitive ability, such as managing financial matters, engaging in shopping activities, and preparing meals [7].

According to the World Health Organization (WHO) survey of 2022, 46.1% of adults 60 years and older worldwide are living with a disability, and this figure is increasing in tandem with the rapid global aging phenomenon [8]. China has one of the highest proportions of older citizens globally, with over 14% of individuals living in China classified as older people (ie, 65 years and older) [8]. This proportion is expected to increase to 30% by 2050, along with a massive drop in the elderly support ratio (defined as the number of people of “working age” [15-64 years] divided by those aged ≥65 years) from 9 in 2010 to 3 in 2050, which is comparable to that of the United States and Germany [9]. A meta-analysis conducted in 2022 reported that the cumulative prevalence of functional disability in China exceeds 30 million [10]. Moreover, the WHO projects that by 2050, the number of older adults living with a functional disability in China will increase to 66 million [11]. Given the high and rising prevalence of functional disabilities among the growing older adult population, exploring the key factors influencing the risk of functional disabilities is crucial to establish appropriate prevention and intervention strategies.

Duration of sleep plays a crucial role in the development of physiological functions that impact health, showing correlations with an increased risk of cognitive decline, depression, cardiovascular diseases, osteoporosis, and stroke [12-17]. However, a consensus has yet to be reached about the association between the duration of sleep and potential risk of functional disability.

In 2016, the National Survey of Midlife Development in the United States presented evidence that insufficient sleep is an independent and important factor contributing to physiological function disability [18]. A cohort study conducted in China with 1798 individuals 90 years or older demonstrated that sleep duration of 8 to 10 hours was associated with the lowest risk of experiencing an ADL disability, whereas a sleep duration exceeding 12 hours was associated with a heightened risk of experiencing ADL disability [19]. However, there is limited knowledge about the correlations of sleep duration with IADL disability. A US study with 136 participants, predominantly comprising older (aged ≥65 years) Black individuals from low-income households, revealed a significant correlation between extended sleep duration surpassing 7.5 hours and worse IADL performance [20]. Similarly, evidence from the National Health Interview Survey spanning 2000 to 2015 showed that an extended sleep duration (≥9 hours) was associated with a higher risk of IADL disability [21]. In addition, owing to interactions with historical, ethnic, economic, and sociocultural factors, the association between sleep duration and functional disability may be heterogeneous across different countries and regions [22,23]. Therefore, the aim of this study was to examine the association between sleep duration and functional disability in China using a nationally representative sample of adults 65 years and older.

Methods

Design and Study Population

This study used data from the China Health and Retirement Longitudinal Study (CHARLS), a comprehensive longitudinal data set designed to provide a representative sample of individuals aged ≥45 years residing in mainland China [3]. The baseline survey of the CHARLS involved a multistage, stratified, probability proportional to sampling method to recruit participants across 150 counties or districts and 450 villages or urban areas throughout the country. Face-to-face interviews were used to obtain the data. A relative or caregiver was asked to complete the survey on behalf of the older adult if assistance was needed. Additional details regarding the CHARLS data set are available elsewhere [3].

The data analyzed in this study were taken from the most recent wave of the CHARLS in 2018, with a sample size of 19,816 participants, to investigate the potential correlation between sleep duration and functional disability. We included only observations without missing values from Chinese older adults aged ≥65 years. The schematic flow of the study sample is depicted in Figure 1. The total sample consisted of 5519 individuals, both with and without functional disability.

This cross-sectional study followed the STROBE (Strengthening the Reporting of Observational Studies in Epidemiology) reporting guideline [24].
Ethical Considerations
The protocols followed in the CHARLS were aligned with the principles of the Declaration of Helsinki [25]. This study obtained ethical approval from the Institutional Review Board at Peking University (IRB00001052-11015, IRB00001052-11014) [3]. All participants in the CHARLS provided written informed consent. The participants received a physical examination report for their participation. The data were deidentified and treated for confidentiality by anonymization and coding.

Sleep Duration
The self-reported data on sleep duration were obtained from a question about the average number of hours participants slept per night over the course of the previous month. Based on previous research experience of the CHARLS on sleep [26,27], we divided respondents into seven sleep duration groups (≤4, 5, 6, 7, 8, 9, and ≥10 hours per night) in the analyses. In addition to assessing sleep duration, we also explored daytime napping habits and established four distinct groups based on napping duration: 0, <30, 30-90, and >90 minutes [28].

Functional Disability
Functional disability was evaluated by asking participants about their ability to perform several domains of routine tasks included in the ADL and IADL questionnaires without special equipment [29]. The response of each item was categorized into four levels: “No, I do not encounter any problem”; “I experience difficulty but am still able to perform the task”; “Yes, I encounter trouble and require assistance”; and “I am unable to perform the task.” According to previous studies [30-32], respondents were recorded as having a functional disability if they reported any challenge in any of the six ADLs (dressing, bathing, eating, getting in and out of bed, using the toilet, and managing urination and defecation) or the five IADLs (household chores, cooking, shopping, paying bills or managing assets, and taking medications) [29,33].

Covariates
The CHARLS structured questionnaire was completed using in-person interviews to gather participants’ demographic characteristics as covariates, including age (years), gender (man or woman), tobacco use (never and current use), alcohol consumption (never and current use), education level (illiterate, primary or middle school, and high school or higher), residential location (urban or rural area), marital status (married or cohabiting and other statuses), afternoon napping habits, chronic
disease status, sampling weights, and depression status. The 10-item Center for Epidemiological Studies Depression Scale was used to differentiate between individuals with and without depression according to a cut-off score of 10 [34-36]. The chronic illness condition of participants was assessed using self-reported noncommunicable diseases (NCDs), including hypertension, diabetes, dyslipidemia, heart issues, stroke, liver diseases, renal diseases, lung diseases, arthritis, and stomach disorders. Based on a recent analysis of CHARLS data on sleep duration [37], individuals were divided into three chronic disease groups: “none” (no NCD), “mild” (1-2 NCDs), or “severe” (more than 3 NCDs). Sampling weights were incorporated based on sets of cross-sectional individual weights that included adjustments for nonresponse by individuals and households.

Statistical Analysis
Participants’ characteristics are summarized as numbers and percentages for categorical variables and as mean and SD values for continuous variables and were divided according to the functional disability status (including ADL and IADL disability). Missing baseline data were handled by a multiple-imputation method, which is a widely used approach to compensate for missing data via generating predictions for each missing value multiple times, resulting in a data set containing no missing values [38]. Multivariable generalized linear models (GLMs) were established using a binomial family and log-link functions to investigate the associations between sleep duration and functional disability status. Restricted cubic-splines analyses with four specific knots located at the 5th, 25th, 75th, and 95th centiles of the exposure distribution were performed to assess dose-response relationships between sleep duration and functional disability. The GLM and restricted cubic-spline models were adjusted for potential confounders, including age, gender, education, marital status, tobacco and alcohol use, afternoon napping, residential location, depression status, chronic disease condition, and sampling weights. All analyses accounting for the complex survey design of the CHARLS were conducted using Stata version 14.0 (Stata Corp). Statistically significant findings were defined by a two-sided \( P \) value below .05.

Results

Sample Characteristics
A total of 5519 participants (2471 men and 3048 women) were included in this analysis with a mean age of 73.67 years, including 2800 (50.73%) participants with a functional disability, 1978 (35.83%) with an ADL disability, and 2299 (41.66%) with an IADL disability. Table 1 provides a summary of the baseline characteristics of the participants according to disability status. The average sleep duration was 6.04 hours for the total sample, with the majority of respondents reporting sleep durations less than 7 hours. Moreover, individuals with functional disability exhibited an average reduction in sleep duration in comparison to that of individuals without functional disability (5.80 vs 6.28 hours). A similar pattern emerged among individuals with ADL and IADL disabilities. Notably, older participants; women; nonsmokers; alcohol abstainers; those who are single; rural inhabitants; as well as those with lower educational attainment, depression, and a higher burden of chronic illnesses had greater rates of functional disability.
Table 1. Baseline characteristics of participants according to functional disability status.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total sample</th>
<th>Functional disability</th>
<th>ADL&lt;sup&gt;a&lt;/sup&gt; disability</th>
<th>IADL&lt;sup&gt;b&lt;/sup&gt; disability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N=5519)</td>
<td>Yes (n=2800)</td>
<td>No (n=2719)</td>
<td>Yes (n=1978)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.80 (2.64)</td>
<td>6.28 (2.12)</td>
<td>5.73 (2.72)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.21 (2.20)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>5.78 (2.71)</td>
<td></td>
<td>5.80 (2.64)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>6.04 (2.41)</td>
</tr>
<tr>
<td>Sleep duration per night (hours), mean (SD)</td>
<td>6.04 (2.41)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours of sleep per night, n (%)</td>
<td></td>
<td>663 (20.59)</td>
<td>755 (32.84)</td>
<td>663 (20.59)</td>
</tr>
<tr>
<td>≤4</td>
<td>1418 (25.69)</td>
<td>900 (32.14)</td>
<td>518 (19.05)</td>
<td>381 (14.68)</td>
</tr>
<tr>
<td>5</td>
<td>792 (14.35)</td>
<td>411 (14.68)</td>
<td>273 (13.80)</td>
<td>109 (5.51)</td>
</tr>
<tr>
<td>6</td>
<td>995 (18.03)</td>
<td>450 (16.07)</td>
<td>302 (15.27)</td>
<td>182 (9.20)</td>
</tr>
<tr>
<td>7</td>
<td>695 (12.59)</td>
<td>259 (9.25)</td>
<td>182 (9.20)</td>
<td>513 (14.49)</td>
</tr>
<tr>
<td>8</td>
<td>885 (16.04)</td>
<td>392 (14.00)</td>
<td>268 (13.55)</td>
<td>206 (5.82)</td>
</tr>
<tr>
<td>9</td>
<td>315 (5.71)</td>
<td>145 (5.18)</td>
<td>109 (5.51)</td>
<td>411 (14.68)</td>
</tr>
<tr>
<td>≥10</td>
<td>419 (7.59)</td>
<td>243 (8.68)</td>
<td>166 (8.39)</td>
<td>214 (9.31)</td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>73.67 (6.41)</td>
<td>74.79 (6.79)</td>
<td>72.53 (5.76)</td>
<td>74.83 (6.87)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>73.03 (6.04)</td>
</tr>
<tr>
<td>Man</td>
<td>2 471 (44.77)</td>
<td>1099 (39.25)</td>
<td>1372 (50.46)</td>
<td>1681 (47.47)</td>
</tr>
<tr>
<td>Woman</td>
<td>3048 (55.23)</td>
<td>1701 (60.75)</td>
<td>1347 (49.54)</td>
<td>1860 (52.53)</td>
</tr>
<tr>
<td>Education, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>1416 (61.59)</td>
</tr>
<tr>
<td>Illiterate</td>
<td>2067 (37.45)</td>
<td>1230 (43.93)</td>
<td>837 (30.78)</td>
<td>1210 (34.17)</td>
</tr>
<tr>
<td>Primary or middle school</td>
<td>3132 (56.75)</td>
<td>1464 (52.29)</td>
<td>1668 (61.35)</td>
<td>2082 (58.80)</td>
</tr>
<tr>
<td>High school or above</td>
<td>320 (5.80)</td>
<td>106 (3.79)</td>
<td>214 (7.87)</td>
<td>71 (3.59)</td>
</tr>
<tr>
<td>Tobacco use&lt;sup&gt;c&lt;/sup&gt;, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>249 (7.03)</td>
</tr>
<tr>
<td>Never</td>
<td>3158 (57.93)</td>
<td>1677 (60.61)</td>
<td>1481 (55.18)</td>
<td>1681 (47.47)</td>
</tr>
<tr>
<td>Current</td>
<td>2293 (42.07)</td>
<td>1099 (39.39)</td>
<td>1203 (44.82)</td>
<td>1860 (52.53)</td>
</tr>
<tr>
<td>Alcohol use, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>1416 (61.59)</td>
</tr>
<tr>
<td>Never</td>
<td>3959 (71.73)</td>
<td>2120 (75.71)</td>
<td>1839 (67.64)</td>
<td>1479 (74.77)</td>
</tr>
<tr>
<td>Current</td>
<td>1560 (28.27)</td>
<td>680 (24.29)</td>
<td>880 (32.36)</td>
<td>1061 (29.96)</td>
</tr>
<tr>
<td>Married or cohabiting, n (%)</td>
<td>3754 (68.02)</td>
<td>1780 (63.57)</td>
<td>1974 (72.60)</td>
<td>2506 (70.77)</td>
</tr>
<tr>
<td>Area of residence&lt;sup&gt;d&lt;/sup&gt;, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>1434 (62.37)</td>
</tr>
<tr>
<td>Rural</td>
<td>4337 (78.65)</td>
<td>2304 (82.34)</td>
<td>2033 (74.85)</td>
<td>2702 (76.39)</td>
</tr>
<tr>
<td>Urban</td>
<td>1177 (21.35)</td>
<td>494 (17.66)</td>
<td>683 (25.15)</td>
<td>835 (23.61)</td>
</tr>
<tr>
<td>Depression&lt;sup&gt;e&lt;/sup&gt;, n (%)</td>
<td></td>
<td>1312 (46.86)</td>
<td>819 (30.12)</td>
<td>1163 (32.84)</td>
</tr>
<tr>
<td>Daytime napping (minutes), n (%)</td>
<td></td>
<td></td>
<td></td>
<td>1083 (47.11)</td>
</tr>
<tr>
<td>None</td>
<td>1972 (35.73)</td>
<td>1016 (36.29)</td>
<td>956 (35.16)</td>
<td>1259 (35.55)</td>
</tr>
<tr>
<td>≤30</td>
<td>417 (9.22)</td>
<td>244 (8.71)</td>
<td>227 (8.35)</td>
<td>304 (8.59)</td>
</tr>
<tr>
<td>31-90</td>
<td>2058 (37.29)</td>
<td>995 (35.54)</td>
<td>1063 (39.10)</td>
<td>1352 (38.18)</td>
</tr>
<tr>
<td>≥90</td>
<td>1018 (18.45)</td>
<td>545 (19.46)</td>
<td>473 (17.40)</td>
<td>626 (17.68)</td>
</tr>
<tr>
<td>Chronic disease condition, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>669 (16.77)</td>
</tr>
<tr>
<td>None</td>
<td>2679 (48.54)</td>
<td>1174 (41.93)</td>
<td>1505 (55.35)</td>
<td>1874 (52.92)</td>
</tr>
<tr>
<td>Mild</td>
<td>2389 (43.29)</td>
<td>1333 (47.61)</td>
<td>1056 (38.84)</td>
<td>945 (41.10)</td>
</tr>
</tbody>
</table>

<sup>a</sup> IADL: instrumental activities of daily living
<sup>b</sup> ADL: basic activities of daily living
<sup>c</sup> Tobacco use: current smoker or former smoker
<sup>d</sup> Area of residence: rural vs. urban
<sup>e</sup> Depression: Mini-Mental State Examination (MMSE) score
IADL

Disability
ADL^a disability
IADL^b disability

Severe

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total sample (N=5519)</th>
<th>Functional disability</th>
<th>ADL^a disability</th>
<th>IADL^b disability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes (n=2800)</td>
<td>No (n=2719)</td>
<td>Yes (n=1978)</td>
<td>No (n=3541)</td>
</tr>
<tr>
<td></td>
<td>255 (11.09)</td>
<td>229 (6.47)</td>
<td>222 (11.22)</td>
<td>229 (6.47)</td>
</tr>
<tr>
<td></td>
<td>No (n=2299)</td>
<td>Yes (n=2800)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>196 (6.09)</td>
<td>255 (11.09)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^aADL: activity of daily living.
^bIADL: instrumental activity of daily living.
^cMissing data for 139 (1.11%) participants.
^dMissing data for 17 (0.14%) participants.
^eDefined as a score of 10 or greater on the 10-item Center for Epidemiologic Studies Depression scale.

Trajectories for Sleep Duration According to Functional Disability Status

The trajectories of sleep duration for groups classified according to different types of functional disability are depicted in Figure 2. The sleep duration of respondents with no functional disability followed a flat curve, whereas there was a U-shaped association between sleep duration and age among respondents with functional disability (Figure 2A). Individuals in the age range of 65 and 75 years with any functional disability had a substantially shorter sleep duration than those 75 years and older (mean 6.11, SD 2.22 vs mean 5.98, SD 2.43). Respondents with functional disability showed an increasing trend in sleep duration after the age of 75 years (mean change per 1 age point later in life of 0.09 hours). Respondents with ADL disability showed a rapid decline in sleep duration between the ages of 65 and 75 years, which increased after the age of 75 years (Figure 2B). Similar patterns are illustrated in Figure 2C, showing that for respondents with IADL disability, the sleep duration trajectories took on a U shape with age, and the shortest sleep duration was evident around the age of 75 years.

Figure 2. Trajectories of sleep duration across individuals according to functional disability status. Graphs display analog values (lines) of the sleep duration with 95% CIs (shaded areas) for any functional disability (A), activity of daily living disability (B), and instrumental activity of daily living disability (C).
Associations Between Sleep Duration and Functional Disability

Table 2 presents the association between sleep duration and functional disability status. In the unadjusted model (model 1), both individuals reporting shorter sleep durations (≤4, 5, and 6 hours) and those reporting longer sleep durations (8, 9, and ≥10 hours) per night exhibited a significantly increased risk of functional disability compared to the reference group (7 hours of sleep/night). The same trend was observed in the relationship between different sleep durations and IADL. Correspondingly, participants reporting shorter sleep durations (≤4 and 5 hours) or longer sleep durations (9 and ≥10 hours) demonstrated a significantly higher odds of ADL disability compared to the reference group.

After adjusting for other potential confounding factors (model 2), the significant associations between shorter sleep durations (≤4, 5, and 6 hours) or longer sleep durations (8, 9, and ≥10 hours) and functional disability persisted, whereas the association of a longer sleep duration (9 hours) to IADL disability dissipated and the association of a longer sleep duration (8 hours) to ADL disability emerged (Table 2).

Table 2. Associations between sleep duration and functional disability status in participants from the 2018 China Health and Retirement Longitudinal Study.

<table>
<thead>
<tr>
<th>Sleep duration per night (hours)</th>
<th>Functional disability</th>
<th>ADL&lt;sup&gt;a&lt;/sup&gt; disability</th>
<th>IADL&lt;sup&gt;b&lt;/sup&gt; disability</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR&lt;sup&gt;c&lt;/sup&gt; (95% CI)</td>
<td>P value</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>Model 1&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤4</td>
<td>2.92 (2.42-3.52)</td>
<td>&lt;.001</td>
<td>2.58 (2.12-3.14)</td>
</tr>
<tr>
<td>5</td>
<td>1.81 (1.48-2.23)</td>
<td>&lt;.001</td>
<td>1.48 (1.19-1.85)</td>
</tr>
<tr>
<td>6</td>
<td>1.39 (1.14-1.69)</td>
<td>.001</td>
<td>1.23 (0.98-1.52)</td>
</tr>
<tr>
<td>7</td>
<td>Reference</td>
<td>—</td>
<td>Reference</td>
</tr>
<tr>
<td>8</td>
<td>1.34 (1.09-1.63)</td>
<td>.005</td>
<td>1.22 (0.98-1.52)</td>
</tr>
<tr>
<td>9</td>
<td>1.43 (1.09-1.88)</td>
<td>.009</td>
<td>1.49 (1.12-1.98)</td>
</tr>
<tr>
<td>≥10</td>
<td>2.32 (1.81-2.97)</td>
<td>&lt;.001</td>
<td>1.85 (1.43-2.39)</td>
</tr>
<tr>
<td>Model 2&lt;sup&gt;f&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤4</td>
<td>2.65 (2.08-3.38)</td>
<td>&lt;.001</td>
<td>2.16 (1.68-2.78)</td>
</tr>
<tr>
<td>5</td>
<td>1.78 (1.34-2.37)</td>
<td>&lt;.001</td>
<td>1.39 (1.04-1.88)</td>
</tr>
<tr>
<td>6</td>
<td>1.67 (1.29-2.15)</td>
<td>&lt;.001</td>
<td>1.26 (0.96-1.66)</td>
</tr>
<tr>
<td>7</td>
<td>Reference</td>
<td>—</td>
<td>Reference</td>
</tr>
<tr>
<td>8</td>
<td>1.51 (1.18-1.94)</td>
<td>.001</td>
<td>1.33 (1.02-1.74)</td>
</tr>
<tr>
<td>9</td>
<td>1.57 (1.12-2.21)</td>
<td>.01</td>
<td>1.66 (1.16-2.38)</td>
</tr>
<tr>
<td>≥10</td>
<td>2.23 (1.60-3.09)</td>
<td>&lt;.001</td>
<td>1.64 (1.17-2.29)</td>
</tr>
</tbody>
</table>

<sup>a</sup>ADL: activity of daily living.

<sup>b</sup>IADL: instrumental activity of daily living.

<sup>c</sup>OR: odds ratio.

<sup>d</sup>Model 1 was unadjusted.

<sup>e</sup>Not applicable.

<sup>f</sup>Model 2 was adjusted for age, gender, education, marital status, tobacco use, alcohol use, afternoon napping, residence, depression, chronic disease status, and sampling weights.

Subgroup Analyses

The findings of subgroup analyses by gender and age are shown in Figure 3 and Figure 4, respectively, where significant effects specific to gender and age were noted. Compared to men, women who slept for shorter durations (≤4, 5, and 6 hours) or longer durations (8, 9, and ≥10 hours) per night were more likely to experience functional disability (Figure 3). Regarding age-specific effects, participants in the older group (75 years and older) who slept for fewer than 6 hours or more than 8 hours were more likely to develop a functional disability than the younger group (74 years and younger) (Figure 4).
Nonlinear Relationship Between Sleep Duration and Functional Disability

Restricted cubic-splines analyses were conducted to visually represent the associations between the duration of sleep and functional disability. We found a U-shaped relationship between sleep duration and functional disability, even after accounting for confounding factors. As shown in Figure 5, the risk of functional disability was negatively correlated with sleep duration until it bottomed out at 7 hours (odds ratio [OR] 0.85, 95% CI 0.79-0.91; \( P < .001 \)). Nevertheless, there was a substantial increase in the risk of functional disability when the duration of sleep exceeded 7 hours (OR 1.16, 95% CI 1.05-1.29; \( P < .001 \)).
Figure 5. Nonlinear relationship of sleep duration and functional disability status. The adjusted odds ratio is presented accounting for potentially confounding factors (green solid line) in the relationship of sleep duration with 95% CIs (dotted lines) for functional disability. The red solid line is the reference line for the association at an odds ratio of 1.0.

Discussion

Principal Results

Our findings obtained from a nationally representative sample of 5519 participants 65 years and older in China indicated that shorter or longer sleep durations are associated with a higher risk of functional disability, including limitations in ADLs and IADLs. In the case of the older population, a minimum risk of 7 hours of sleep is associated with a reduced likelihood of experiencing functional disability. Based on previous studies highlighting functional disability as a notable risk factor for sleep disorders in the older population [39], the findings of this study suggest that the association between sleep and functional disability can exhibit a bidirectional nature.

Limitations

Our study has several limitations. First, the data of sleep duration and functional disability assessment were collected via self-report by participants, which could be subject to information bias. Second, the sample for this study comprised individuals from China who were 65 years of age or older. This particular demographic characteristic may limit the generalizability of the findings to different age cohorts, geographical areas, or ethnic backgrounds. Third, this study adopted a cross-sectional design, which limits the ability to show a causal association. It is plausible to consider that older adults with functional disability may require prolonged periods of sleep and a reduced duration of sleep could potentially serve as an initial indication of dysfunction. Therefore, it is necessary to conduct further comprehensive cohort studies to validate these findings.

Comparison With Prior Work

Functional ability refers to an individual’s capacity to engage in daily living and social activities according to their own intentions and preference. However, it is discouraging that the prevalence of functional disability was determined to be 41% across the entire sample in China [40]. The aging process is commonly accompanied by disturbances in sleep patterns, which have been linked to notable alterations in brain function and a decline in overall quality of life [41]. Moreover, various population-based studies [42-44], including both cross-sectional and longitudinal study designs, have demonstrated that the prevalence of cognitive decline may be linked to an increased risk of experiencing functional disability. Epidemiological research has also documented a U-shaped association between sleep duration and cognitive decline, indicating a significant trend in the association between sleep duration and functional disability [45-47]. Nevertheless, a definitive consensus has not yet been attained and there is still a dearth of studies examining the association between sleep duration and functional disability.

Despite previous investigations into the association between sleep duration and self-care function, the majority of these studies have primarily concentrated on a single form of functional disability. One study involving nightshift workers in the United States found that individuals with a shorter sleep duration (<7 hours/day) had the highest prevalence of sleep problems (61.8%) and the highest prevalence of an impaired ADL score (24.8%) [48], revealing a correlation between shorter sleep duration and higher risk of ADL disability. Another study focusing on patients with dementia discovered that a longer sleep duration was associated with ADL disability [49].
Similarly, a cross-sectional study that recruited 1798 participants older than 90 years found that a long sleep duration (≥12 hours) may be associated with an increased risk of ADL disability for this population [19]. Peng et al [50] found that after accounting for potential confounders such as age and gender, both longer and shorter sleep durations were linked to a heightened risk of IADL disability, which was consistent with the results of this study. Furthermore, instead of solely focusing on the relationship between shorter or longer sleep duration and functional disability among different sample sets, we simultaneously explored the impact of both shorter and longer sleep durations on functional disability within the same sample group. Additionally, concerning functional disability, we separately explored the associations between sleep duration and ADL and IADL disabilities, demonstrating that both shorter and longer sleep durations have an influence on functional health among the older population.

Implications and Contribution
The mechanism that accounts for the association between sleep duration and functional disability has yet to be fully explained; however, several hypotheses have been proposed. Luo et al [51] demonstrated that inadequate and excessive sleep durations were associated with an increased likelihood of hypertension among Chinese individuals. The increased susceptibility to cardiovascular illnesses associated with this elevated risk can have a direct influence on the overall well-being and functional abilities of older adults. Another recent study revealed that both insufficient and excessive sleep durations are associated with an increased risk of late-life dementia, a condition that significantly impairs the ability of older individuals to perform their daily tasks [52]. Other relevant studies indicated that sleep was correlated with changes in epigenetic mechanisms such as DNA methylation and histone modifications, which can also lead to cognitive dysfunctions such as learning and memory disruption [53-55]. This evidence has provided a new avenue for exploring the mechanism underlying the relationship between sleep duration and functional disability. Interventions aimed at good sleep hygiene may have the capacity to yield favorable outcomes in terms of improving physiological function among older adults. As evidenced by the empirical findings of this study, maintenance of the recommended sleep duration (7 hours) might play a crucial role in the health of older adults.

Our subgroup analyses suggested that women and older adults aged ≥75 years with shorter sleep durations (≤4, 5, and 6 hours) or longer sleep durations (8, 9, and ≥10 hours) per night were more likely to experience functional disability. Gender differences in sleep duration among the older population cannot be ignored. With changes in biological life cycles and the extreme hormonal fluctuations occurring with advancing age, women are at an increased risk for sleep disturbances (including insomnia and hypersomnia) [56]. Further, the sleep disturbances occurring during menopause can be an independent risk factor associated with arterial stiffness, which can result in a higher incidence of osteoarthritis that is in turn highly related to dysfunction [57]. With respect to age-specific effects, changes in sleep duration are a part of the normal aging process and also may enhance cellular aging in the later years of life [58]. According to the findings from an umbrella review [59], extreme sleep durations (including shorter and longer sleep durations) were more likely to be associated with an elevated risk of all-cause mortality, cognitive disorders, and type 2 diabetes in the general population. The circadian oscillations that alter body functions, including sleep, become less pronounced in older age, which can increase the risk of functional disability [60]. Collectively, these results emphasize the importance of addressing the complex needs of the population experiencing functional disabilities, particularly among women and older adults.

Conclusions
In conclusion, more attention should be paid to older individuals with shorter or longer sleep durations than recommended (7 hours). The precise mechanisms underlying the association between sleep duration and functional disability in the older population require further investigation.

Acknowledgments
We thank the China Health and Retirement Longitudinal Study (CHARLS) team for providing data and all the respondents for their contribution. This work was supported by a Beijing University of Chinese Medicine Evidence-based Chinese Medicine Research and Development Fund phase I grant (90020172220020). The funder played no role in the design, methods, subject recruitment, data collection, analysis, or preparation of the article.

Data Availability
The data sets presented in this study can be found in online repositories of the China Health and Retirement Longitudinal Study [61]. The data sets generated and analyzed during this study are available from the corresponding author on reasonable request.

Authors’ Contributions
HR and YF had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. ML, YD, BF, and XZ contributed to the hypothesis and study design and interpreted the results. HL and CL analyzed the data. ML wrote the manuscript. HR and YF revised the manuscript. All authors contributed to the article and approved the submitted version.
References


https://aging.jmir.org/2024/1/e53548


Abbreviations

ADL: activity of daily living
CHARLS: China Health and Retirement Longitudinal Study
GLM: generalized linear model
IADL: instrumental activity of daily living
NCD: noncommunicable disease
OR: odds ratio
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
WHO: World Health Organization
Identifying Predictive Risk Factors for Future Cognitive Impairment Among Chinese Older Adults: Longitudinal Prediction Study

Collin Sakal¹, BA, MSc; Tingyou Li¹; Juan Li², BA, MA, PhD; Xinyue Li¹, BA, MS, PhD

¹ Corresponding Author: Xinyue Li, BA, MS, PhD

Abstract

Background: The societal burden of cognitive impairment in China has prompted researchers to develop clinical prediction models aimed at making risk assessments that enable preventative interventions. However, it is unclear what types of risk factors best predict future cognitive impairment, if known risk factors make equally accurate predictions across different socioeconomic groups, and if existing prediction models are equally accurate across different subpopulations.

Objective: This paper aimed to identify which domain of health information best predicts future cognitive impairment among Chinese older adults and to examine if discrepancies exist in predictive ability across different population subsets.

Methods: Using data from the Chinese Longitudinal Healthy Longevity Survey, we quantified the ability of demographics, instrumental activities of daily living, activities of daily living, cognitive tests, social factors and hobbies, psychological factors, diet, exercise and sleep, chronic diseases, and 3 recently published logistic regression–based prediction models to predict 3-year risk of cognitive impairment in the general Chinese population and among male, female, rural-dwelling, urban-dwelling, educated, and not formally educated older adults. Predictive ability was quantified using the area under the receiver operating characteristic curve (AUC) and sensitivity-specificity curves through 20 repeats of 10-fold cross-validation.

Results: A total of 4047 participants were included in the study, of which 337 (8.3%) developed cognitive impairment 3 years after baseline data collection. The risk factor groups with the best predictive ability in the general population were demographics (AUC 0.78, 95% CI 0.77-0.78), cognitive tests (AUC 0.72, 95% CI 0.72-0.73), and instrumental activities of daily living (AUC 0.71, 95% CI 0.70-0.71). Demographics, cognitive tests, instrumental activities of daily living, and all 3 recreated prediction models had significantly higher AUCs when making predictions among female older adults compared to male older adults and among older adults with no formal education compared to those with some education.

Conclusions: This study suggests that demographics, cognitive tests, and instrumental activities of daily living are the most useful risk factors for predicting future cognitive impairment among Chinese older adults. However, the most predictive risk factors and existing models have lower predictive power among male, urban-dwelling, and educated older adults. More efforts are needed to ensure that equally accurate risk assessments can be conducted across different socioeconomic groups in China.

(JMIR Aging 2024;7:e53240) doi:10.2196/53240

KEYWORDS

- cognitive impairment
- China
- prediction
- predictions
- predict
- predictor
- predictors
- risk
- risks
- population
- demographic
demographics
- gerontology
- geriatric
- geriatrics
- older adult
- older adults
- elder
- elderly
- older person
- older people
- ageing
- aging
- MCI
- cognitive
- cognition
- machine learning
- variable
- variables
- model
- models
- mild cognitive impairment

Introduction

China’s aging population has led to cognitive impairment becoming increasingly burdensome to society [1,2]. In 2020, more than 68 million Chinese older adults had mild cognitive impairment, dementia, or Alzheimer disease [3]. The economic and social burden of cognitive impairment has led to calls for improving risk assessments and prioritizing early diagnoses [1,3]. Given China’s limited number of geriatric psychiatrists, researchers have turned to developing prediction models to identify older adults at risk of cognitive impairment for preventative interventions [4-8]. However, no study has compared the predictive ability of known risk factors side by side, and our understanding of which factors are the most useful for developing prediction models is limited. Furthermore, population characteristics vary widely across China, but it is unknown which risk factors are the most predictive in different socioeconomic groups, and existing prediction models have primarily been tested in the general population alone. To understand how to best predict future cognitive impairment and to develop more targeted prediction models for population subgroups, the predictive ability of known risk factors and...
A plethora of modifiable and nonmodifiable risk factors that are associated with cognitive impairment among Chinese older adults have been identified. Previous studies have found that increased age, limited functional independence, alcohol consumption, hypertension, and depression are significantly associated with cognitive impairment [3,9-12]. Protective factors have also been identified, namely, good sleep quality, sleeping sufficiently for many hours per night, exercise, and increased social participation [3,9,13-17]. It is also known that the prevalence of cognitive impairment in China differs across population subsets such as male and female individuals, rural and urban dwellers, and older adults with different levels of education [3]. Explanations for such discrepancies include different social patterns and literacy rates across the sexes and across regions with varying degrees of rurality, rates of depression in rural areas, and levels of education [3]. More years of education has also been associated with a greater cognitive reserve, which protects against future impairment [3]. Nevertheless, it is unclear which risk factors are the most useful for predicting future cognitive impairment across different subpopulations in China. Although we have a reasonable understanding of which characteristics make a person more likely to develop cognitive impairment, our understanding of which parts of a person’s health profile most accurately predict their risk of developing cognitive impairment is limited.

In addition to known risk factors, it is unclear if existing prediction models for future cognitive impairment are equally accurate across different socioeconomic groups in China. Several published models have reported areas under the receiver operating characteristic curve (AUCs) greater than 0.80 in their development cohorts [5,7,8], but each model has only been tested on the general population. Nearly all existing models make predictions by leveraging measures of cognition, age, and education. Additional covariates vary from model to model and include factors such as instrumental activities of daily living (IADL), hobbies such as gardening and watching television, marital status, and others. Examining the predictive ability of existing models across population subsets would allow us to identify where more efforts are needed to improve risk assessments for cognitive impairment, further our understanding of which subpopulations are more difficult to conduct risk assessments within, and provide a more thorough evaluation of existing prediction models than has been reported previously.

In this study, we quantified the ability of 9 risk factor groups and 3 existing models to predict future cognitive impairment among Chinese older adults. We examined how well demographics, IADL, activities of daily living (ADL), cognitive tests, social factors and hobbies, psychological factors, diet, exercise and sleep, chronic diseases, and 3 recently published models predict future cognitive impairment in the general population and among male, female, rural-dwelling, urban-dwelling, educated, and not formally educated older adults. To our knowledge, this study is the first to comprehensively compare the ability of known risk factors to predict future cognitive impairment and the first seeking to identify which subsets of the Chinese population need greater attention to improve the accuracy of risk assessments.

**Methods**

**Data Source and Study Design**

The Chinese Longitudinal Healthy Longevity Survey (CLHLS) is a prospective cohort study of Chinese older adults that contains information on demographics, cognitive function, lifestyle factors, chronic diseases, and more [18,19]. The CLHLS began in 1998, and follow-up surveys have been conducted every 2-3 years since. The data include older adults from 23 of China’s provinces that together make up 85% of the country’s total population.

We used the 2011 and 2014 CLHLS waves in our study. Baseline characteristics were gathered from the 2011 survey and used to predict if an individual became cognitively impaired by 2014. CLHLS participants younger than the age of 60 years or those with cognitive impairment at baseline were excluded. Sample size calculations were conducted following the methodology for multivariable prediction models by Riley et al [20]. This study is presented following the TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) guidelines where appropriate [21,22].

**Ethical Considerations**

The CLHLS received ethics approval from the Duke University Institutional Review Board (Pro000062871) and Peking University’s Biomedical Ethics Committee (IRB00001052–13,074). Written informed consent was given by all participants prior to the survey interviews. This study secondarily analyzed anonymized data from the CLHLS.

**Measuring Cognitive Function**

Cognition was assessed through the Chinese-language version of the Mini Mental State Examination (MMSE) [23]. MMSE scores range from 0 to 30, with lower scores indicating worse cognitive function. We adopted education-specific cutoffs that have been previously validated in the Chinese older adult population to indicate cognitive impairment [24]. Those with no formal education and MMSE scores less than 18 were labeled as cognitively impaired, as were those with 1-6 years of education with scores less than 21 and those with more than 6 years of education with scores less than 25 [24].

**Risk Factor Groups**

**Overview**

A total of 9 groups containing known risk factors for cognitive impairment were considered in this study: demographics, ADL, IADL, cognitive tests, social factors and hobbies, psychological factors, exercise and sleep, diet, and chronic diseases. The risk factor groups were chosen by selecting parts of a person’s health profile previously found to be associated with developing cognitive impairment [9]. Each group is briefly described below, and a complete list of the variables in each group can be found in Multimedia Appendix 1.
Demographics Group
The demographics group contained each individual’s age, sex, years of education, household income, marital status, and residence location (city, town, or rural area).

ADL Group
The ADL group included each person’s ability to bathe, get dressed, use the toilet, get in and out of bed, control urination and bowel movements, and eat food.

IADL Group
The IADL group covered tasks that require thinking, organizational, and physical independence. The IADL group included an older adult’s ability to visit neighbors, go shopping, cook, wash clothes, walk continuously for 1 km, lift a bag of groceries, crouch and stand up, and take public transportation.

Cognitive Tests Group
The cognitive tests group included scores from each subsection of the MMSE: orientation, naming, registration, calculation, attention, recall, and language. Scores from each section were included as separate variables.

Social Factors and Hobbies Group
The social factors and hobbies group included whether a person grows vegetables, gardens, reads newspapers and books, looks after pets or animals, plays cards or mahjong, and participates in social activities.

Psychological Factors Group
The psychological factors group included the following factors that primarily relate to depression and anxiety: whether a person is generally optimistic, keeps their belongings organized, is generally anxious, is often lonely, makes decisions independently, feels useless with age, was happier when they were younger, and felt sad for more than 2 consecutive weeks over the past year.

Exercise and Sleep Group
The exercise and sleep group included whether someone currently exercises, whether they used to exercise, as well as the self-reported duration and quality of sleep.

Diet Group
The diet group contained information on each person’s staple food; if they eat fresh fruits and vegetables; the main flavor of the dishes they cook; how frequently they consume meat, fish, eggs, sugar, and tea; if they consume alcohol; the type of alcohol they consume; and the frequency of alcohol consumption.

Chronic Diseases Group
The chronic diseases group included the presence or absence of hypertension, diabetes, heart disease, blood disease, and cardiovascular disease.

Recreating Existing Prediction Models
Prediction models were selected based on the following criteria: the model was developed for use in China, was reproducible using the CLHLS, and had an AUC of >0.75 during development. We selected 3 models published in Zhou et al [8], Hu et al [5], and Wang et al [7]. Each model was developed for use in the general Chinese population and showed excellent predictive performance (AUC>0.80) during development. All the models we recreated were based on logistic regression, which returns predictions by summing weighted values of each covariate before the sum is passed through the logistic function to produce predicted probabilities between 0 and 1. The logistic regression model recreated from Zhou et al [8] included age, a functional independence score based on ADL, baseline MMSE score, chewing ability, visual function, history of stroke, whether the participant watches TV or listens to the radio, and whether the participant grows flowers or raises pets. From Hu et al [5], the recreated model included age, marital status, IADL, and baseline MMSE score. Lastly, the model from Wang et al [7] included age; education; sex; ADL; baseline MMSE; and whether the participant gardens, reads newspapers or books, plays mahjong or cards, watches TV, or listens to the radio.

Statistical Analysis
All analyses were performed using the R Statistical Software (version 4.0.5; R Foundation for Statistical Computing), and all code required to reproduce the analyses presented herein can be found on the web [25]. Predictive ability was quantified using AUC, sensitivity, and specificity. We assessed the predictive ability of each risk factor group using logistic regression models evaluated through 20 repeats of 10-fold cross validation, which has been recommended to obtain optimism-corrected performance metrics for prediction models [26]. This resulted in 200 training sets and 200 validation sets. All “I don’t know” or “refused” responses in the CLHLSs were set to missing, ordinal variables were integer encoded, and nonordinal categorical variables were dummy encoded. Missing values were imputed on each training and validation set separately using k-nearest neighbors imputation [27]. During each iteration of cross-validation, the data were split into training and validation sets before 9 logistic regression models, each containing all covariates in 1 particular risk factor group, were fit to the training data. Thereafter, each model was used to make predictions on the validation set for the general population and 6 subpopulations: male, female, rural-dwelling, urban-dwelling, educated, and not formally educated older adults. The same procedure was also followed for evaluating the prediction models from Zhou et al [8], Hu et al [5], and Wang et al [7]. Average AUCs and accompanying 95% CIs were calculated across the 200 validation-set AUCs for each model in this study. Sensitivity and specificity curves, 1 from each validation set, were also plotted for the risk factor group models.

Results
Given a binary outcome, a population-level prevalence of 0.20, a conservatively estimated Cox-Snell $R^2$ of 0.09, and 24 predictors in the largest risk factor group, the sample size required for this study was determined to be 1065 with 213 events to minimize the risk of overfitting, reduce the chance of overly optimistic performance metrics, and ensure that the models have sufficient data to estimate the overall risk of cognitive impairment in our sample. After excluding CLHLS participants with cognitive impairment at baseline and those...
younger than the age of 60 years, a cohort of 4047 Chinese older adults were included, of which 337 (8.3%) developed cognitive impairment. The average age of the cohort was 79.8 (SD 9.4) years, and 2037 (50%) were male. The group that developed cognitive impairment was older at baseline (89.1 vs 79.0 years) with a lower average baseline MMSE score (25.1 vs 27.7). A full description of the cohort’s characteristics can be found in Table 1, and the distribution of variables in each risk factor group can be found in Multimedia Appendix 2.

### Table 1. Baseline cohort characteristics.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>All participants (N=4047)</th>
<th>Developed cognitive impairment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Yes (n=337)</td>
</tr>
<tr>
<td>Age (y), mean (SD)</td>
<td>79.8 (9.4)</td>
<td>89.1 (9.8)</td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>2037 (50.3)</td>
<td>130 (38.6)</td>
</tr>
<tr>
<td>Female</td>
<td>2010 (49.7)</td>
<td>207 (61.4)</td>
</tr>
<tr>
<td>Years of schooling, mean (SD)</td>
<td>2.8 (3.7)</td>
<td>1.8 (3.2)</td>
</tr>
<tr>
<td>Household income (CN ¥; CN ¥1=US $0.14), mean (SD)</td>
<td>24,483.8 (25,778.6)</td>
<td>22,942.1 (23,198.5)</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married and living with spouse</td>
<td>2033 (50.3)</td>
<td>83 (24.6)</td>
</tr>
<tr>
<td>Married but not living with spouse</td>
<td>89 (2.2)</td>
<td>4 (1.2)</td>
</tr>
<tr>
<td>Divorced</td>
<td>8 (0.2)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Widowed</td>
<td>1862 (46.1)</td>
<td>246 (73)</td>
</tr>
<tr>
<td>Never married</td>
<td>46 (1.1)</td>
<td>4 (1.2)</td>
</tr>
<tr>
<td>Residential status, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>City</td>
<td>665 (16.4)</td>
<td>57 (16.9)</td>
</tr>
<tr>
<td>Town</td>
<td>1241 (30.7)</td>
<td>89 (26.4)</td>
</tr>
<tr>
<td>Rural area</td>
<td>2141 (52.9)</td>
<td>191 (56.7)</td>
</tr>
<tr>
<td>Baseline MMSE&lt;sup&gt;a&lt;/sup&gt; score, mean (SD)</td>
<td>27.5 (2.8)</td>
<td>25.1 (3.6)</td>
</tr>
<tr>
<td>Follow-up MMSE score, mean (SD)</td>
<td>26.2 (5.2)</td>
<td>12.8 (6.1)</td>
</tr>
</tbody>
</table>

<sup>a</sup>MMSE: Mini Mental State Examination.

As shown in Figure 1A and Table 2, demographics had the best predictive ability in the general population (AUC 0.78, 95% CI 0.77-0.78), followed by cognitive tests (AUC 0.72, 95% CI 0.72-0.73) and IADL (AUC 0.71, 95% CI 0.70-0.71). Social factors and hobbies had a moderate predictive ability (AUC 0.67, 95% CI 0.66-0.68), whereas diet, psychological factors, exercise and sleep, ADL, and chronic diseases all had average AUCs less than 0.60. Demographics, cognitive tests, and IADL also had the best sensitivity and specificity tradeoffs, as shown in Figure 2. By contrast, the sensitivity and specificity curves for the chronic diseases group showed that such risk factors only sometimes resulted in better-than-random predictions.
Figure 1. Average AUC by predictor group and target population. ADL: activities of daily living; AUC: area under the receiver operating characteristic curve; IADL: instrumental activities of daily living.
Table. Predictive ability by target population.

<table>
<thead>
<tr>
<th>Model</th>
<th>Target population, AUC&lt;sup&gt;a&lt;/sup&gt; (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>General population</td>
</tr>
<tr>
<td>Demographics</td>
<td>0.78 (0.77-0.78)</td>
</tr>
<tr>
<td>Cognitive tests</td>
<td>0.72 (0.72-0.73)</td>
</tr>
<tr>
<td>IADL&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.71 (0.70-0.71)</td>
</tr>
<tr>
<td>Social factors and hobbies</td>
<td>0.67 (0.66-0.68)</td>
</tr>
<tr>
<td>Diet</td>
<td>0.59 (0.58-0.60)</td>
</tr>
<tr>
<td>Psychological factors</td>
<td>0.58 (0.57-0.59)</td>
</tr>
<tr>
<td>Exercise and sleep</td>
<td>0.57 (0.57-0.58)</td>
</tr>
<tr>
<td>ADL&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.57 (0.56-0.57)</td>
</tr>
<tr>
<td>Chronic diseases</td>
<td>0.53 (0.52-0.53)</td>
</tr>
<tr>
<td>Wang et al [7]</td>
<td>0.80 (0.80-0.81)</td>
</tr>
<tr>
<td>Zhou et al [12]</td>
<td>0.80 (0.80-0.81)</td>
</tr>
<tr>
<td>Hu et al [5]</td>
<td>0.80 (0.80-0.81)</td>
</tr>
</tbody>
</table>

<sup>a</sup>AUC: area under the receiver operating characteristic curve.

<sup>b</sup>IADL: instrumental activities of daily living.

<sup>c</sup>ADL: activities of daily living.
Figure 1D shows that most risk factor groups had significantly higher AUCs when making predictions among older adults with no formal education compared to those with some education. The only exceptions were the ADL and exercise and sleep groups. Among those with no formal education, demographics, cognitive tests, and IADL had AUCs of 0.79 (95% CI 0.78-0.79), 0.72 (95% CI 0.72-0.73), and 0.71 (95% CI 0.70-0.71), respectively. When making predictions among those with some education, demographics, cognitive tests, and IADL had average AUCs of 0.73 (95% CI 0.72-0.74), 0.64 (95% CI 0.63-0.66), and 0.64 (95% CI 0.62-0.65), respectively.

The existing prediction models recreated in this study all had good predictive ability in the general population. Each model had an average AUC of 0.80 (95% CI 0.80-0.80.1). However, each model had significantly higher AUCs when making predictions in female individuals compared to male individuals, in rural dwellers compared to urban dwellers, and in those with no formal education compared to those with some education. Complete results can be found in Table 2 and Figure 3. The TRIPOD checklist for this study can be found in Checklist 1.

Demographics had a significantly better predictive ability when making predictions among rural dwellers (AUC 0.80, 95% CI 0.79-0.81) compared to urban dwellers (AUC 0.74, 95% CI 0.73-0.75). Similarly, IADL showed a higher average AUC among rural dwellers (AUC 0.73, 95% CI 0.72-0.73) compared to urban dwellers (AUC 0.68, 95% CI 0.67-0.69). As shown in
Figure 1C and Table 2, significantly higher AUCs among rural dwellers were also observed for the diet, psychological factors, and chronic diseases groups.

Demographics, cognitive tests, and IADL also had the highest average AUCs when making predictions for male and female individuals, although the predictive ability varied between the 2 sexes. The demographics group had a higher average AUC when making predictions in female individuals compared to male individuals (0.81, 95% CI 0.80-0.82 vs 0.72, 95% CI 0.71-0.73), as did the IADL group (0.72, 95% CI 0.71-0.73 vs 0.66, 95% CI 0.65-0.67) and the cognitive tests group (0.72, 95% CI 0.71-0.73 vs 0.70, 95% CI 0.69-0.71). The dietary group had a significantly higher AUC when making predictions among male individuals (0.61, 95% CI 0.60-0.62) compared to female individuals (0.57, 95% CI 0.56-0.58). No significant differences were observed for the social factors and hobbies group, and all other remaining groups has AUCs less than 0.60 for both male and female individuals. Full results can be found in Figure 1B and Table 2.

Figure 3. Predictive ability of existing models [5,7,8]. AUC: area under the receiver operating characteristic curve.
Discussion

Principal Findings

In this study, we quantified the ability of 9 risk factor groups and 3 prediction models to predict future cognitive impairment in the general Chinese population and 6 population subsets: male, female, rural-dwelling, urban-dwelling, educated, and not formally educated older adults. In the general population, the risk factor groups with the best predictive ability were demographics (AUC 0.78, 95% CI 0.77-0.78), cognitive tests (AUC 0.72, 95% CI 0.72-0.73), and IADL (AUC 0.71, 95% CI 0.70-0.71). The most predictive risk factors and the existing models performed inconsistently across socioeconomic groups and had significantly higher AUCs when making predictions for female individuals and older adults with no formal education compared to male individuals and older adults with some education.

Our study showed that the 3 existing prediction models had significantly lower AUCs when predicting future cognitive impairment among male, urban-dwelling, and educated Chinese older adults compared to female, rural-dwelling, and not formally educated older adults. Despite the only shared risk factors in all 3 models being age and baseline MMSE score, significant differences in predictive ability were consistent across every model. One explanation is that risk factor differences between those who developed cognitive impairment and those who did not were larger among the groups where more accurate predictions were made. For example, the difference in average age between female older adults who did and did not become cognitively impaired was 11.8 years, whereas for male older adults, it was 7.0 years. Similarly, among those with no formal education, the difference in baseline MMSE score between those with and without cognitive impairment at follow-up was 2.67 compared to 1.45 among those with some education. In addition, the prevalence of cognitive impairment in our sample was higher among female, rural-dwelling, and not formally educated older adults, meaning that the models had more events to learn from. Indeed, previous studies using nationally representative data have also reported higher prevalence estimates among these groups [3]. Our results indicate that targeted prediction models for specific socioeconomic groups are needed in China to make equally accurate risk assessments across sex, residential status, and education level. Several studies have called for such models [28,29], but as of this writing, none have been developed in China.

Out of the 9 risk factor groups, we found that demographics, cognitive tests, and IADL best predict future cognitive impairment in the general Chinese population and across sex, residential status, and education level. Demographics are often included in prediction models for cognitive impairment [28,30-32], and we suggest that they continue to be leveraged because of their predictive power and ease to collect. Associations between chronic diseases, ADL, psychological factors, and diet with cognitive impairment among Chinese older adults have been established [13,33-39], but such factors showed moderate predictive ability in our study. To our knowledge, dietary factors have not been incorporated into existing prediction models in China, but they had higher AUCs than commonly used risk factors such as psychological factors, ADL, and chronic diseases. In fact, chronic diseases did not make significantly better than random predictions among male (AUC 0.50, 95% CI 0.49-0.51), urban-dwelling (AUC 0.50, 95% CI 0.50-0.51), and not formally educated (AUC 0.51, 95% CI 0.49-0.52) older adults. Hence, in addition to providing a ranking of the most predictive risk factor groups, our study is the first to show that dietary factors warrant consideration when predicting future cognitive impairment among Chinese older adults.

Many risk factor groups had significantly different AUCs across population subsets. Similar to the existing models we recreated, our study revealed that the most predictive risk factors (demographics, cognitive tests, and IADL) had significantly higher AUCs when making predictions among female and not formally educated older adults compared to male and educated older adults. As was the case with the recreated models, this likely resulted from the distributions of risk factors being more separable between those who developed cognitive impairment and those who did not in the groups where more accurate predictions were made. Given the lack of available evidence, it is unclear whether the discrepancies in predictive ability found in our study generalize outside of China, and future work may seek to perform similar analyses elsewhere.

Limitations

Our study has several limitations. The source code was not available for the models we selected to recreate in this study, but we explicitly followed all preprocessing, variable derivation, and model creation procedures described in the original papers during the model replication process. The AUCs of each model in the general population in this study were consistent with the reported AUCs in the original papers, suggesting that the models were properly recreated from scratch. To facilitate future research, we have further made our code publicly available. Second, the CLHLS is not nationally representative, although it does include older adults from 23 of China’s provinces. The exercise and sleep group did not include objective measurements of physical activity and sleep. Self-reported exercise and sleep are often inaccurate, and we suggest that the results be interpreted with caution for the exercise and sleep group. Lastly, the data used in this study were from 2011 to 2014. Future studies may wish to collect new data and further validate the results presented herein.

Conclusions

Out of the 9 risk factor groups, our study found that demographics, cognitive tests, and IADL best predicted future cognitive impairment among Chinese older adults and had significantly better predictive ability among female and not formally educated older adults compared to male and educated older adults. Similarly, every existing model we recreated made significantly better predictions among female, rural-dwelling, and not formally educated older adults. Our study suggests that more targeted predictions models for cognitive impairment are needed to make equally accurate risk assessments across different socioeconomic groups in China and provides...
foundational evidence that can support variable selection for such models.

Acknowledgments

This project was funded by the City University of Hong Kong, Hong Kong SAR, China, internal research grant #9610473. The funder had no role in the design, analyses, or interpretations of this study. The authors would like to thank Peking University for providing the data.

Authors’ Contributions

CS, XL, and JL designed the study. CS and TL performed the statistical analyses. CS and XL wrote the manuscript with additional input from JL.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Predictors in each risk factor group and prediction model.

[DOCX File, 21 KB - aging_v7i1e53240_app1.docx]

Multimedia Appendix 2

Cohort characteristics for each covariate from every risk factor group.

[DOCX File, 65 KB - aging_v7i1e53240_app2.docx]

Checklist 1

TRIPOD (Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis) checklist.

[DOCX File, 26 KB - aging_v7i1e53240_app3.docx]

References


Abbreviations

ADL: activities of daily living
AUC: area under the receiver operating characteristic curve
CLHLS: Chinese Longitudinal Healthy Longevity Survey
IADL: instrumental activities of daily living
MMSE: Mini Mental State Examination
TRIPOD: Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis

©Collin Sakal, Tingyou Li, Juan Li, Xinyue Li. Originally published in JMIR Aging (https://aging.jmir.org), 22.3.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Three Perspectives on Older Adults’ Daily Performance, Health, and Technology Use During COVID-19: Focus Group Study

Ortal Cohen Elimelech1, MSc; Sara Rosenblum1, PhD; Michal Tsadok-Cohen1, MSc; Sonya Meyer2, PhD; Simona Ferrante3, PhD; Naor Demeter4, PhD

1 Corresponding Author: Ortal Cohen Elimelech, MSc

Abstract

Background: During COVID-19 lockdowns, older adults’ engagement in daily activities was severely affected, causing negative physical and mental health implications. Technology flourished as a means of performing daily activities in this complex situation; however, older adults often struggled to effectively use these opportunities. Despite the important role of older adults’ social environments—including their families and health professionals—in influencing their technology use, research into their unique perspectives is lacking.

Objective: This study aimed to explore the daily activity performance, health, and technology use experiences of healthy independent Israeli adults (aged ≥65 years) during COVID-19 from a 3-dimensional perspective: older adults, older adults’ family members, and health professionals.

Methods: Nine online focus groups, averaging 6-7 participants per group, were conducted with older adults, family members, and health professionals (N=59). Data were analyzed using thematic analysis and constant comparative methods.

Results: The intertwining of daily activity performance and health emerged as a central theme, with differences between the groups. Older adults prioritized their self-fulfilling routines based on motivation and choice, especially in social-familial activities. In contrast, family members and health professionals focused on serious physical and mental health COVID-19–related consequences. A consensus among all three groups revealed the meaningful role of technology use during this period in bridging functional limitations. Participants delved into technology’s transformative power, focusing on the need for technology to get engaged in daily activities.

Conclusions: This study illustrates the profound interplay between daily activity performances, physical and mental health, and technology use, using a 3-dimensional approach. Its focus on technology’s uses and benefits sheds light on what older adults need to increase their technology use. Interventions for improving digital activity performance can be tailored to meet older adults’ needs and preferences by focusing on motivational and preference-related activities.

(JMIR Aging 2024;7:e53141) doi:10.2196/53141

KEYWORDS
daily activity; health; technology use; older adult; qualitative study; focus group; COVID-19

Introduction

Background

Occupational scientist Ann Wilcock [1] claimed that “occupation and health are inseparable,” meaning that a person’s ability to function and perform in daily activities and occupations, especially meaningful ones, directly affects their health [2]. Previous studies have emphasized the substantial role of daily activity performance in maintaining and enhancing older adults’ health. More information should be gathered to explore the unique daily experiences of older adults during a crisis such as the COVID-19 pandemic, since they cope with specific age-related changes, including physical, psychological, and social factors [3].

Although COVID-19 presented additional major challenges to older adults’ daily activity performance and health, some remained active, most notably by using technology [4,5]. Since technology has become integral to modern life, it is essential to understand technology use among older adults in terms of their daily activity performance and health [6]. One major factor affecting older adults’ success in daily functioning and adopting technology is their social environment [7,8], including family and health professionals [9]. For example, family members can support technology challenges (choosing suitable devices and
introducing basic functions), and health professionals can suggest appropriate technology [8]. Thus, a deep understanding of daily activity performance and technology use necessitates consideration of family members’ and health professionals’ experiences.

Nevertheless, previous studies focused almost entirely on older adults’ perspectives regarding daily activity performance. For example, Israeli studies of activity performance during the COVID-19 pandemic focused on adults (18 years and older) in general without addressing the specific needs and experiences of older adults (65 years and older) [5,10]. It remains unclear how older adults adjusted their daily activity performance in the face of the COVID-19 outbreak. Similarly, a study conducted among family members and health professionals regarding technology use did not specifically address the needs of older adults in terms of using technology to promote health in times of crisis [11]. Therefore, it is necessary to gain greater insight into Israeli older adults’ experiences of their daily activity performance and technology use during COVID-19 through the various perspectives of older adults, family members, and gerontology health professionals. In this study, technology is considered with daily activity performance. Hence, the term technology refers to devices that adults already own (eg, mobile phones, computers, and tablets) for carrying out these activities.

Prior Studies

Older Adults’ Daily Activity Performance and Health During the COVID-19 Pandemic

Globally, people are living longer, with more options to discover new interests, such as education or professional training. Alternatively, they might contribute positively to their families and society by caring for their grandchildren or volunteering for various charitable activities. Despite this empowering perspective, a central aspect—health—strongly influences their ability to engage in daily activities [12].

Extraordinary circumstances such as the COVID-19 pandemic present a substantial concern for older individuals’ health. Worldwide, they reported restrictions on daily activities, such as banking, shopping, hiking [13], and leisure activities [14]. These changes negatively affected their physical and mental health [15,16]. Hence, exploring older adults’ daily experiences during the COVID-19 pandemic can provide insight into their needs.

The occupational perspective, which provides an in-depth view of human doing, highlights the contribution of daily activity performance to health. This perspective further emphasizes the impact of the activities’ context, time, and role in doing, being, becoming, and belonging [17]. Situations like COVID-19 may negatively affect older adults’ ability to function (doing) and their sense of self (being). It can present obstacles to achieving future goals (becoming) and particularly affect older adults’ ability to participate socially (belonging) [18]. The occupational perspective can demonstrate how humanity adapted to COVID-19 changes by adapting activities to what was available at the time.

Studies in numerous global regions showed that older adults adjusted to the COVID-19–related activity changes in various ways. They may have acquired a new skill or knowledge (eg, learning a foreign language), modified the nature of their activities, or adjusted the time allocated to each activity. Thus, they might have engaged in social, leisure, physical, and educational activities throughout the pandemic—but in new ways. Maintaining routines and participating in such meaningful activities regularly facilitated individuals maintaining their mental health [19,20].

Israeli Society’s Distinctive Characteristics

Previous cross-sectional studies among Israeli adults (18 years and older) during COVID-19 highlighted the importance of maintaining daily routines. Their findings showed that Israeli adults discontinued some activities and modified their environment, often opting for solitude or staying home [5,10]. Cultural influences greatly determine how older adults navigate daily activities and use technology [21], thus requiring insight into their experiences. Despite the earlier studies, Israeli older adults’ experiences of their performance in daily activities during crises still require clarification.

Technology-Supported Daily Activity Performance During the COVID-19 Pandemic

Previous studies focused mainly on older adults’ self-reports on their technology use [4,19-26]. However, because social support, often provided by family members or health professionals, can overcome gaps between technology and barriers to using it [8], it is also vital to explore the person’s social environment. The studies that were conducted among family members and health professionals focused on issues like assistive technology [27], robots allowing aging in place [28], technology for care services [11], strategies and barriers for communication technologies [29], and accepting personal alerting devices [30]. None directly addressed the older adults’ specific needs for using technology in meaningful daily activities to promote health in crisis periods.

This Study’s Goal

This study aimed to fill the literature gaps by exploring daily activity performance and technology use experiences among healthy, independent, older Israeli adults during COVID-19 using a 3-dimensional perspective (including older adults, their family members, and health professionals).

Methods

Overview

This study is part of a larger project, Empathic Platform to Personally Monitor, Stimulate, Enrich, and Assist Elders and Children in Their Environment (ESSENCE). The ESSENCE used opportunities arising during the COVID-19 pandemic to study vulnerable populations, including older adults and children. This paper provides the results of focus groups exploring older adults’ daily activity performance experiences with technology in the COVID-19 pandemic context. The focus groups provided a multi-perspective window through diverse
daily communication, allowing us to learn more about human experiences, desires, and concerns.

**Ethics Approval**
This study was approved by the University of Haifa Faculty Ethics Committee (086/21).

**Procedures**

**Recruitment**
We recruited participants using snowball sampling and social media. Participants received information about the study, signed online consent forms, and completed a short demographic questionnaire. We honored their willingness to share their experiences with gift certificates.

**Data Collection**
The data were collected between February 2021 and July 2022 (Israel was in lockdown part of this time). Using three types of focus groups (older adults, family members, and health professionals) allowed for triangulation, strengthening the results’ validity [31].

The 9 focus group sessions were conducted by the first and third authors. We calculated the focus group size (6-7 participants each) according to Krueger and Casey’s [32] guidelines. The moderators created a relaxed and friendly environment, encouraging participant interaction and continued data collection until the main issues were repeated and theoretical saturation was reached [33].

**Participants**
The 9 focus groups comprised 3 clusters. The clusters were:

1. Older adults (65 years and older) living independently at home and able to use a computer and Zoom software (n=20)
2. Family members of older adults deeply familiarized with the older adults’ routines (n=19)
3. Health professionals with at least 5 years of experience working with older adults (n=20)

Table 1 presents each group’s sociodemographic characteristics.
Table. Sociodemographic characteristics of study participants.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Older adults (n=20)</th>
<th>Family members (n=19)</th>
<th>Health professionals (n=20)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>10 (50)</td>
<td>16 (84)</td>
<td>18 (90)</td>
</tr>
<tr>
<td>Male</td>
<td>10 (50)</td>
<td>3 (16)</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Level of family closeness, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>First-degree family links</td>
<td>—</td>
<td>15 (78)</td>
<td>—</td>
</tr>
<tr>
<td>Second-degree family links</td>
<td>—</td>
<td>4 (21)</td>
<td>—</td>
</tr>
<tr>
<td>Health profession, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Doctor</td>
<td>—</td>
<td>—</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Nurse</td>
<td>—</td>
<td>—</td>
<td>4 (20)</td>
</tr>
<tr>
<td>Occupational therapy</td>
<td>—</td>
<td>—</td>
<td>6 (30)</td>
</tr>
<tr>
<td>Physiotherapy</td>
<td>—</td>
<td>—</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Social worker</td>
<td>—</td>
<td>—</td>
<td>3 (15)</td>
</tr>
<tr>
<td>Day care director</td>
<td>—</td>
<td>—</td>
<td>2 (10)</td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>70.8 (3.81)</td>
<td>39.16 (10.52)</td>
<td>49.85 (11.75)</td>
</tr>
<tr>
<td>Range</td>
<td>66-80</td>
<td>21-55</td>
<td>33-70</td>
</tr>
<tr>
<td>Education (years), mean (SD)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>15.7 (2.45)</td>
<td>16.5 (3.29)</td>
<td>—</td>
</tr>
<tr>
<td>Range</td>
<td>—</td>
<td>—</td>
<td>5-35</td>
</tr>
<tr>
<td>Experience working with older adults (years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>—</td>
<td>—</td>
<td>17.95 (9.49)</td>
</tr>
<tr>
<td>Range</td>
<td>—</td>
<td>—</td>
<td>5-35</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>0 (0)</td>
<td>3 (16)</td>
<td>—</td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>3 (15)</td>
<td>4 (21)</td>
<td>—</td>
</tr>
<tr>
<td>Married</td>
<td>14 (70)</td>
<td>12 (63)</td>
<td>—</td>
</tr>
<tr>
<td>Widowed</td>
<td>3 (15)</td>
<td>0 (0)</td>
<td>—</td>
</tr>
<tr>
<td>Employment status, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full-time</td>
<td>1 (5)</td>
<td>10 (53)</td>
<td>—</td>
</tr>
<tr>
<td>Part-time</td>
<td>2 (10)</td>
<td>9 (47)</td>
<td>—</td>
</tr>
<tr>
<td>Retired</td>
<td>17 (85)</td>
<td>0 (0)</td>
<td>—</td>
</tr>
</tbody>
</table>

*Not applicable.

Each focus group lasted approximately 60-70 minutes. They were conducted via Zoom videoconferencing software, video- and audio-recorded in Hebrew, and stored on the researchers’ password-protected computers. The recordings were transcribed without details that could identify the participants. Once transcribed, the original records were deleted to maintain anonymity and confidentiality.

The researchers wrote reflective comments capturing verbal and nonverbal interactions during each focus group. Because the groups were conducted online, particular attention was paid to the older participants’ needs. For instance, we offered support for difficulties in using technology. Support included help setting up an internet connection in another room and repeating the conversation issues. The broad perspective on older adults’ use of technology during the focus group further enhanced the study’s credibility.

**Research Tools**

The moderator’s questions were customized for each group. However, the session structure, developed based on findings in the relevant literature [34], was similar for all three focus group types. It was designed to create a relaxing environment where participants felt free to share their experiences and thoughts. We invited the participants to introduce themselves and asked a general opening question to encourage active participation in the discussion: “What is your most enjoyable activity in your free time?” Additional questions were asked related to older adults’ general daily function and routines, for example, “Would you be able to summarize your ordinary day, starting from the...
moment you wake up to the moment you go to sleep.” daily activity performance during the COVID-19 period (“How has the COVID-19 affected your daily activity performance in comparison to the routine before the COVID-19?”), and technology (“Which technology devices do you use and how do you use them?” and “How has the use of technology changed since the COVID-19?”).

Data Analysis

The data were analyzed according to thematic analysis and the constant comparative method [35] using Word (Microsoft Corporation) and Excel (Microsoft Corporation) worksheets. The researchers repeatedly reread the transcripts and reflective comments to familiarize themselves with the data and used the Office software to create memos related to the text. The data were then sorted, highlighted, and categorized by cases, and comparisons were made. Comprehensive coding was conducted to produce themes, and selective coding was performed to fit the theme precisely. Last, the quotations were rearranged into new categories and translated into English. We paid particular attention to the similarities and differences between participants’ experiences in each group and between the three groups [33,34,36,37]. In-depth discussions of any disagreements among the researchers were held in a series of research meetings with a researcher in the field and a qualitative researcher [38], improving the findings’ trustworthiness [36].

Results

Overview

We identified two main themes, with two subthemes each:

- Daily activity performance and health are intertwined.
- Changes in daily activity performance affect health.
- Meaningful activities shape a healthy routine.
- Technology use bridges functional limitations.
- “It forced them to use technology.”
- Opportunities to engage in daily activities.

Whereas older adults described how they adapted during the pandemic and used it for self-fulfillment, the family members and health professionals discussed its devastating effect, especially relative to physical and mental health.

Data Analysis

The data were analyzed according to thematic analysis and the constant comparative method [35] using Word (Microsoft Corporation) and Excel (Microsoft Corporation) worksheets. The researchers repeatedly reread the transcripts and reflective comments to familiarize themselves with the data and used the Office software to create memos related to the text. The data were then sorted, highlighted, and categorized by cases, and comparisons were made. Comprehensive coding was conducted to produce themes, and selective coding was performed to fit the theme precisely. Last, the quotations were rearranged into new categories and translated into English. We paid particular attention to the similarities and differences between participants’ experiences in each group and between the three groups [33,34,36,37]. In-depth discussions of any disagreements among the researchers were held in a series of research meetings with a researcher in the field and a qualitative researcher [38], improving the findings’ trustworthiness [36].

Daily Activity Performance and Health Are Intertwined

Changes in Daily Activity Performance Affect Health

Miriam (a 69-year-old female participant) illustrated the older adults’ changing daily activity performance during the lockdown period when they were required to stay home:

“My previous routine included the whole culture thing, And I miss it very much...I live in a private house in Israel, and we saw the grandchildren. We were careful...We were restricted from going for a walk. Now, we can do a little more. It is very lacking, and it greatly affects my mood that I can’t go to the theater or plays or concerts. It is a central part of my life.

Although she met with others, Miriam’s inability to participate in cultural activities, which she considered vital, affected her mood.

Racheli exemplified the family members’ descriptions of emotional reactions, changes in daily activity performance, and concerns about their older family members:

The Corona interrupted [my mother’s] blossom and somewhat stole the joy of her retirement...It was also very frustrating for her...It seems we cared more about her than she cared about herself...And suddenly, all the classes and everything moved to Zoom...It was also hard for her to follow, so [she lost] all her fun from “brain strength” or [other] kinds of things...She was keeping herself, feeling that she was suddenly deteriorating, and she struggled terribly with it—with this feeling and with trying not to be a burden on us, not expressing her loneliness so that we didn’t feel it, too.

Racheli portrayed her mother’s frustration during this crisis period, in which activities were reduced, compared to her earlier delight in retirement. Although the situation and her loneliness affected Racheli’s mother, she did not share her feelings with the family.

The professional therapists’ perspectives added to this dimension while foregrounding the health-related implications. One physiotherapist referred to medical treatment, stating, “I don’t think we’ve even talked about all the medical treatments that people forgo in order not to go to, I don’t know what, hospitals, even dental treatments.”

A family therapist expounded, “I think the last year brought death closer to older people in a very present way...and they are very busy with their physical health and fears that were not there before.”

The professional therapists included physical health alongside the fear, exemplifying how this crisis affected both physical and mental health and how they intertwine. A geriatrician described how the mental state caused stressful reactions:

I saw terrible examples of loneliness and anxiety...In the beginning, it really was all of us, even the children who, every day before Corona, would come in and hug their parents, and parents, as usual, would get along. Here, they came to the door of the house because they said it was forbidden to enter and it was forbidden to meet, and they left the food for them—left it for them. Rung the bell, but it wasn’t enough. I had terrible examples of people who were afraid, all the time using alcohol gel, and came [to me] with wounds on their hands and all kinds of terrible anxiety.
Meaningful Activities Shape a Healthy Routine

Lea’s (66-year-old woman) words illustrated how the older adults described adaptations that helped them cope with challenges in the COVID-19 period:

I also had a difficult time when everything was closed. In the more serious lockdowns, the children refused to come. I told them, “Come, I’m right at the limit in terms of age,” but they didn’t agree. On Passover, they came and put something for me by the door. And they told me, “We left something for you by the door.” Flower...It was really...wow. On the other hand, I had a dog, so I would go for walks. So it helped because the dog took me for a walk in a way, so that was good. The garden saved me. The garden and the dog, and that’s it; some phones.

Daily meaningful outdoor activities such as walking her dog and tending her garden supported Lea during this challenging period. Similarly, Rebecca (a 70-year-old woman) demonstrated adaptation to the new situation to the point where she enjoyed it:

Compared to before? Obviously, I haven’t left the house since the end of February. I was at home. And I’m telling you the truth, I really enjoyed it. I picked up, arranged, got to a lot of things that I hadn’t gotten to in years...I adapted myself to the thought that this is what there is.

The health professionals’ focus groups discussed reasons for the shift in activities and their meaning. A paramedic described:

Until now, they were really assisted. First, their day would have been filled with grandchildren who would come,...you know, there would be a reason to go to the kitchen and make the meatballs this child loves...There was this thing, now, because of the whole last year that was not created, so they simply look for the meaning in other forms.

According to that paramedic, these circumstances forced older adults to find solutions to their difficulties and continuity in routine, such as preparing meals for grandchildren, which supported their coping.

Technology Use Bridges Functional Limitations

“It Forced Them to Use Technology”

The COVID-19 pandemic restrictions required individual adjustments in daily activity performance. One prominent change was how technology was used in daily life. Michel (a 67-year-old man) described needing technology to perform social and leisure activities:

I don’t know how I would have gotten through this period without technology. I started taking a course that stopped due to the Coronavirus. But during all these periods, there are, of course, Zooms and all kinds of lectures, even WhatsApp conversations between groups of friends.

Rebecca described keeping her life active with technology while learning new skills:

This (virus) is unknown...It has stopped the whole world, and I just wait for it to pass and learn how to behave afterward. I was very active every day. And even now in Corona, I participate in Zoom, learning new things. There is nothing to be done; you must realize that this is the situation, and that is what there is.

She described COVID-19 as a period of inactivity but also emphasized the need to acquire new skills, especially those related to technology.

The family members’ and health professionals’ groups also demonstrated how technology use became a necessity during COVID-19. Olivia, the daughter of an older adult, explained:

My father retired a short time before the Coronavirus, and then the Coronavirus started, and it really worried us; it was huge. And we said, “Okay, even this [retirement] transition, which seemed difficult for him, and then Corona.” And it seemed way too much. But surprisingly—maybe because it was something worldly—my father really found himself. Like maybe because there was no alternative, and everyone was now in some kind of madness, so he really surprised us for the better. As if he really found himself. He is not a technological person, but he really found himself in Zooms, one after the other, such as lectures and all. Even to the extent that they hang up on us because they now have some Zoom.

Olivia’s quote shows that even though her father apparently had no knowledge of technology before the COVID-19 pandemic, he learned to use it because it was necessary and there was no alternative. The restrictions and lockdowns decreased communications with the outside world. A gerontology social worker noted technology use as a communication issue to consider:

We knew this even before the Coronavirus. There was a lot of work in this area, but obviously, during the Coronavirus, [digital literacy] jumped by hundreds of percent. And it was a sudden realization that it’s something almost existential,...that it’s nice fun, and it’s enriching. You need a real existence for connection with the world.

A nurse added, “It forced them to use [technology] because, otherwise, there is no option to communicate with the world.” The social worker and the nurse emphasized that technology use becomes a virtual requirement for not only engaging in meaningful activities but also staying connected with the outside world.

Opportunities to Engage in Daily Activities

The older adult participants depicted their technology use according to their desires for meaningful activities, divided into three main domains: instrumental activities of daily living (IADL), leisure, and social activities.

Aharon (a 66-year-old man) described how COVID-19 affected IADL, especially shopping:
A notable thing that changed is a funny thing: Yes, shopping in the supermarket has become the whole of the Corona only through the Internet. It...yes, continues with it, almost never physically visiting the store.

Jacob (a 79-year-old man) described technology use for leisure activities:

We tried to compensate (which was not possible in the pre-Corona period) during the Corona period...We play bridge online; there is an online option to play bridge with opponents from all over the world who enter the table and open and play for an hour or so...In the game of bridge, both partners play but do not see the cards each has. We arranged the two computers, so we sit back-to-back and don’t see the partner’s cards, and that’s how we play, and it’s very nice. Sometimes, when you lose, you get quite frustrated, but that’s the game.

Although Jacob played the game on an online platform, it seemed to provide the same experience as regular gaming. It made him feel he was compensating for activities he could not perform.

The availability of technology seemed to provide older adults more opportunities to participate in lessons than they were used to. Shira (69-year-old woman) described:

The timing of Corona was good for me. I felt I was already very tired when this rest came to me, and what helped me with Corona was Zoom. I am alone at home, and I have a son who lives with his partner. It helped me overcome this period. From morning to evening, on Zoom all the time. You don’t have to sit in front of the screen; you also listen to the lectures.

Shira described the benefits of the COVID-19 period, bringing her opportunities to rest and interact using Zoom.

Emma, a daughter of an older adult, added:

Now, during the Corona period, we bought [my mother] a smartphone that she didn’t have before, and this made it very accessible for us to share with her...Every time she would send pictures by email, and when everyone is on WhatsApp and presses a button and forwards to everyone, then an email is already something a little more like sending a letter by post with a stamp. And she didn’t always understand: “What’s so hard for you? Well, send.” The phone was a gift from us. We bought it for her birthday, but it gave us the option to share with her more and gave her another tool to help her orient herself.

Communicating through advanced technology apparently provided Emma’s mother opportunities to not only stay in touch with family but also feel a sense of belonging.

The health professionals described opportunities for older adults’ remote health care. A paramedic explained:

They receive remote medical treatments...like they were sitting in front of doctors once upon a time. Who thought of it? So, coming from a generation connected with a wire to the wall, like a curled phone, everything was so terrible when talking face-to-face. Suddenly, they learned the whole young generation of today, and they are inside it. And I say up to higher studies, which is beautiful...They know how to use a computer at such a level as the Open University.

The paramedic exemplified the incredible changes in the older population receiving health care during COVID-19, adjusting their habits to the realities of the COVID-19 period when people were required or encouraged to stay home.

Discussion

Principal Findings

COVID-19 had a substantial impact on people’s health. Technology use supported the daily lives of Israeli older adults (65 years and older) and affected their health during the pandemic. Findings based on a 3-dimensional perspective (older adults, family members, and health professionals) highlight the overlaps and differences between perspectives and provide explanations of these findings. Overall, the findings point to older adults’ need to stay active as a factor motivating their technology use.

Older Adults: Meaningful Activities Shape Healthy Routines

The participants described changes in daily activity performance, emphasizing COVID-19 constraints as causing emotional difficulties. Their descriptions demonstrate the effects of time and context on their activities and how doing and being interrelate from an occupational perspective [17]. Older Japanese adults described similar pain and fear emotions regarding their daily activity limitations [39]. The aging process alone includes substantial changes in routine (like retirement) and in the ability to engage in some activities due to biological, psychological, and social changes [3,40]. These aging characteristics, magnified in a crisis, underscore the importance of offering emotional support to older adults with daily activity performance limitations resulting from varying causes.

Despite limitations preventing them from reaching their aspirations and becoming, older adults seek adaptions and opportunities to remain active. They become involved in new activities, gain new skills, and find alternative methods (like using technology) to carry out their daily tasks. In line with previous research [19,20], this study’s participants describe how staying engaged in leisure, educational, and outdoor activities (eg, gardening and walking) supported their daily routines during COVID-19.

Interestingly, the older adult participants tend to emphasize engaging in secondary routines, associated with preference and motivation (eg, leisure and social activities) rather than primary routines. Primary routines are behaviors vital for survival (eg, sleep and personal hygiene) and establish the general pattern of daily life [41,42]. The participants stressed the benefits of meaningful activities of pleasure, social interaction, and mental health, which motivate them to participate [43].
Specifically, older Israeli adults frequently mentioned social activities with family members, including religious rituals. The activities perceived as meaningful vary across cultures [21]. Older adults in Brazil, Italy, Portugal, and the United States viewed their social environment, including family and friends, and religious rituals as coping mechanisms. They underlined the crucial role of the family as a support system during the COVID-19 pandemic [13]. Somewhat differently, the older Israeli adults in this study often referred to activities with their families, indicating that maintaining contact with their relatives was paramount and composed a substantial part of their daily routines. Even with cultural differences, the value of the family to older adults in times of crisis is cross-cultural.

**Family Members: Close Social Environments and Caring for a Loved One**

Like the older participants, the family members described changes they noticed in their daily activity performance, including leisure, educational, and social activities. However, they exhibited greater concern for the older members of their families, prioritizing their loved ones’ safety and emotional support. For instance, they worried about the older adults’ increased vulnerability to the virus, the potential consequences to their health [7], and particularly their loneliness and unwillingness to share their struggles. Witnessing their loved ones’ health deterioration and accelerated aging while experiencing their own COVID-19 challenges could increase the family members’ mental burdens. Since they are a source of support for older adults [7], professional involvement is warranted in supporting both older adults and their family members during times of crisis.

**Health Professionals: Health Implications**

Similarly, to older adults and family members, health professionals outlined changes in meaningful activities. However, they added a main focus on ensuring the health and safety of older adults, drawing upon their expertise and knowledge [44]. Notably, the professional participants represented a wide range of fields; several were older than 60 years, allowing for valuable and unique insights. They assessed the situation accurately and objectively from a medical perspective, being equipped with knowledge of the aging process, and addressed health consequences, including mental and physical implications. They particularly noted loneliness, anxiety, and physical health symptoms associated with COVID-19 in older adults and stressed the importance of maintaining meaningful activities to enhance older adults’ health. Consideration of these factors can contribute to older adults’ health and well-being whenever they are faced with similar situations.

**Focus Group Consensus: Technology Use Promotes Health**

All focus groups agreed on technology’s necessity. The need for technology during the COVID-19 pandemic might have played a role in the rise of older adults adopting and using technology [21]. Technology is indispensable for motivating them and meeting their needs in routine and crisis periods. During the COVID-19 pandemic, older adults had to cope with an unknown situation requiring them to use technology. Adapting to unfamiliar situations and using technology effectively require higher-order cognitive abilities of executive functions (eg, inhibition, working memory, cognitive flexibility, planning, and problem-solving) [45]. Although possibly the first affected by the cognitive dysfunction associated with aging, executive functions have a remarkable ability to maintain physical and mental health [46,47]. Therefore, it is vital to leverage the COVID-19 period by creating opportunities for older individuals to engage in unfamiliar digital activities, which could benefit executive functions and improve physical and mental health.

Technology benefits include opportunities for daily activities, especially IADL (eg, shopping and receiving remote health care), education, leisure, and social activities. As in previous research, this study’s participants described how engaging in digital activities fostered a sense of belonging [48]. Focusing on technology’s uses and benefits sheds light on older adults’ needs, allowing them to increase their use of technology.

**Limitations and Future Research**

It is important to acknowledge this study’s limitations. First, participation in online focus groups requires high functional abilities. Thus, the study might present a limited perspective on digital activity performance because potential participants with less technology proficiency were not included. Further, most participants in all groups were female, possibly introducing gender bias and limiting the diversity of perspectives represented in the study. Finally, the snowball sampling method increased the likelihood that some participants, particularly health professionals, had preexisting relationships.

Future studies should include participants with a more diverse range of technological proficiency and gender. Because participating in digital technology is vital to physical and mental health among the older adult population, further research should be conducted in both a qualitative and quantitative manner exploring technology’s downsides, including its disadvantages and challenges for older adults. This could identify key principles for interventions to promote older adults’ technology use.

**Conclusion**

This study illustrates the profound interplay between daily activities, physical and mental health, and technology use among Israeli older adults (65 years and older) using a 3-dimensional approach. Specifically, it delves into the perspectives of older adults, their family members, and health professionals. In light of an occupational perspective, older adults emphasize family connection activities as a significant aspect of their lives. Thus, when activities are restricted, it is imperative to provide emotional support to them and their families. All three groups emphasize the importance of digital activities for coping with changes in routines and activities and promoting emotional and physical health during crises. These findings apply to various circumstances that older adults may encounter, ranging from health conditions to crises. Whether facing personal challenges such as the loss of a loved one or coping with social upheavals
like war, their daily lives may be affected. The focus on technological uses and benefits during COVID-19 sheds light on what older adults need to increase their technology use. Interventions for improving digital activity performance can be tailored to meet their needs and preferences by focusing on their secondary routines.

**Acknowledgments**

This study has received funding from the European Union’s Horizon 2020 research and innovation program under the call SC1-PHE-CORONAVIRUS-2020-2B (grant agreement 101016112).

**Authors’ Contributions**

OCE contributed to data collection and analysis and writing the original draft. SR contributed to project administration, conceptualization, methodology, validation, analysis, writing, reviewing, and editing. MTC contributed to data collection and analysis. SM contributed to conceptualization, data collection and analysis, validation, writing, reviewing, editing, and visualization. ND contributed toward conceptualization, data collection and analysis, validation, writing, reviewing, editing, and visualization.

**Conflicts of Interest**

None declared.

**References**


27. van Heek J, Ziefle M, Himmel S. Caregivers’ perspectives on ambient assisted living technologies in professional care contexts. Presented at: 4th International Conference on Information and Communication Technologies for Ageing Well and e-Health; Mar 22 to 23, 2018; Funchal, Madeira, Portugal. [doi: 10.5220/0006691400370048]


Abbreviations

**ESSENCE**: Empathic Platform to Personally Monitor, Stimulate, Enrich, and Assist Elders and Children in Their Environment

**IADL**: instrumental activities of daily living

© Ortal Cohen Elimelech, Sara Rosenblum, Michal Tsadok-Cohen, Sonya Meyer, Simona Ferrante, Naor Demeter. Originally published in JMIR Aging (https://aging.jmir.org), 30.5.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Technology Use Among Older Adults and Their Caregivers: Cross-Sectional Survey Study

Shinduk Lee¹, DrPH; Marcia G Ory²,³, PhD; Deborah Vollmer Dahlke³,⁴, DrPH; Matthew Lee Smith³,⁵, PhD

Corresponding Author: Marcia G Ory, PhD

Abstract

Background: Informal caregivers are called upon to provide substantial care, but more needs to be known about technology use among older adult and caregiver dyads.

Objective: This study described technology use among older adults and their caregivers, explored potential correlates of technology use, and highlighted implications for practice.

Methods: A cross-sectional survey was conducted among unpaid caregivers of older adults (n=486). Primary outcomes were self-reported technology (devices and functions) use among caregivers and their oldest care recipient. The concordance of technology use among caregivers and care recipients was also examined. Multivariable regression models were conducted separately for caregivers and care recipients.

Results: Greater proportions of caregivers used all examined technologies, except for the medication alerts or tracking function, than care recipients. Caregivers used an average of 3.4 devices and 4.2 functions, compared to 1.8 devices and 1.6 functions used by their care recipients. Among caregivers, younger age, higher income, and higher education were associated with more technology use (P<.05). Among care recipients, younger age, not having cognitive dysfunction, and caregiver’s technology use were associated with more technology use (P<.05).

Conclusions: Understanding technology use patterns and device adoption across diverse caregiver and care recipient populations is increasingly important for enhancing geriatric care. Findings can guide recommendations about appropriate technology interventions and help providers communicate and share information more effectively with patients and their caregivers.

(JMIR Aging 2024;7:e50759) doi:10.2196/50759

KEYWORDS

technology; caregiving; social determinants of health; health disparities; disparity; disparities; caregiver; caregivers; carers; technology use; usage; gerontology; geriatric; geriatrics; older adult; older adults; elder; elderly; older person; older people; ageing; aging; cross-sectional; survey; surveys; computer use; device; devices; adoption; dyad; dyads

Introduction

Aging is occurring in most parts of the world [1]. Driven by the large baby boomer generation, the American population is aging rapidly [2], with 1 in 4 Americans estimated to be 65 years or older by 2060 [3]. Although healthy aging may be the “new normal” for some [4], normative age-related decline in physical and cognitive functions exists and often results in the need for assistance with daily household and self-care activities. Among 53 million American adult informal caregivers, nearly 42 million cared for adults aged 50 years or older [4,5]. Despite the many hours devoted to informal caregiving [4,5], many older adults face adverse consequences of unmet needs for assistance with daily activities [6,7]. In addition to inadequate caregiving resources for older Americans, the negative impacts of caregiving on caregivers’ health and quality of life have raised significant programmatic and policy concerns [8-10].

An array of innovative technology solutions exists to support older adults’ health, independence, and quality of life, enabling them to age in place [11-14]. These technologies also support caregivers and enhance caregiving for older adults in areas such as fall alert notifications, social supports, communication, and medication scheduling [15,16]. However, older age is frequently considered a prominent factor associated with diminished interest in, and adoption of, technology [17-19]. This can be attributed to unique barriers associated with older age, including the lack of experiences or familiarity and declining physical, cognitive, and sensory functions [20,21]. Although recent data
suggest a narrowing of the age-based gap in the digital divide [5,22,23], data also show that older adults may not use technology to its full potential [5].

The latest national surveillance data showed that about 24% of baby boomers provided informal care [24], with older caregivers tending to provide care for care recipients at similar or older ages [25]. The existence of any age-related difficulty in accessing or using technology is relevant for understanding health technology use among older adults in need of care, as well as their older caregivers who can use technology for caregiving. Although abundant literature discusses technology use among older adults or caregivers [5,25-27], few studies have examined technology use among both older adults and their caregivers [28,29] and the potential relationship between older adults and their caregivers’ technology use [16].

The relationship between older adults and their caregivers is interdependent, extending to their use of technology. Knowing the pattern of technology use in older adults and their caregivers can inform the development of technology-based interventions that are accessible and usable to the aging community. For example, a qualitative study involving patients with type 1 diabetes and their spouses revealed that continuous glucose monitoring technology can enhance spousal engagement in diabetes care, yet it may also introduce sources of tension within the relationship [30]. The qualitative study provided preliminary data to guide the development of a technology-based intervention, called SHARE plus [31]. In another recent study, Shih et al [32] focused on the different types of digital devices and categories of smartphone functions used by caregivers and care recipients compared to those with no caring roles. Shih et al [32] developed a health-related smartphone app for older adults and their caregivers, and their recent work was conducted to improve the design of their smartphone app.

Our study sought to further our understanding of how the use of technology by caregiver and care recipient dyads can guide intervention outlets (eg, digital platforms) and support efficient deployment of the technology interventions, including expanded access and use of technology functions. Therefore, technology use was broadly defined in this study to enable research into the use of diverse types of devices and functions. Devices and functions represent different aspects of technology use. Devices encompass the equipment or hardware of technology, whereas functions pertain to the specific tasks one can perform using technology. For instance, an individual may possess a smartphone, tablet, and computer but only use them for internet browsing. In contrast, another person with access solely to a smartphone may use it for various functions such as email and texting, internet browsing, web-based banking, and more. Analyzing devices and functions independently can offer more precise insights to inform future technology-based interventions for the aging community.

Our key aims were (1) to describe the use of various technologies (ie, both devices and functions) for caregivers and older adult care recipients; (2) to compare technology use among caregivers and older adult care recipients; and (3) to examine potential correlates of caregivers’ and older adult care recipients’ technology use. A study by Lindeman et al [33] offers a conceptual framework for identifying and addressing the challenges in technology-enabled solutions for family caregivers. While not yet a theoretical basis to study and analyze caregivers’ and care recipients’ use of technology, the conceptual framework by Lindeman et al [33] pointed to several factors that influence the caregivers’ technology adoption. The individual-level moderators involved user capacity and family, and socioeconomic moderators encompassed race and ethnicity, income, and geographic location [33]. Rather than directly assessing user capacity, we explored factors potentially linked to user capacity, such as age and education for caregivers, and cognitive dysfunction for care recipients.

Methods

Data Source and Study Population

This study collected cross-sectional data from a web-based survey about technology use among paid and unpaid caregivers of older adults who were recruited through a Qualtrics panel. To be eligible to participate in the web-based survey, the respondents must have provided 8 or more hours of weekly care for at least 1 adult care recipient aged 50 years or older (N=626). Recognizing that caregivers may be caring for multiple individuals, the caregiver was asked to respond to the survey questions in the context of the oldest person to whom they provided at least 8 hours of care. Quota sampling was used to ensure data were collected from a diverse sample reflecting the general characteristics of the US caregiving population.

Predetermined targets were set: 75% of the recruited sample were to be female, 50% at least 50 years old, and no more than 60% White. Geographic targets were also set to represent the regional population proportion (ie, 17.2% in the Northwest, 20.9% in the Midwest, 23.8% in the West, and 38.1% in the South [34]). This study focused on unpaid caregivers (n=486). The web-based survey commenced with a set of screening questions to identify eligible individuals. Those who were not screened out were provided with study information necessary for informed consent. Only those who agreed to participate were invited to complete the web-based survey.

This study differentiated between technology devices and technology functions. The respondents were asked whether they used each of the 7 devices (ie, cell phone, smartphone, tablet, computer, e-reader, voice-activated assistant, and wearable or smartwatch for activity tracking) and 8 functions (ie, communication, ride-sharing, online shopping, online banking, navigation, online entertainment, medication alerts or tracking, and physical activity tracking). The types of devices and functions were determined based on the 2020 AARP Tech and the 50+ Survey report [25]. We dropped some of the minimally used devices (eg, virtual reality device, 1%) and combined functions (eg, instead of individually assessing games, music, and video or movie streaming, they were consolidated into the “online entertainment” category) [25]. The respondents were also asked about their oldest care recipients’ use of the same devices and functions. The total numbers of devices (ranging from 0 to 7) and functions (ranging from 0 to 8) used were calculated separately for caregivers and care recipients.
Socioeconomic and demographic characteristics of the respondents were collected using the web-based survey: age (years), sex (male or female), race and ethnicity (Non-Hispanic White or others), household income (less than US $50,000 or US $50,000 or more), and education (high school graduate and lower educational attainment or higher). The web-based survey also asked about the respondents’ place of residence (zip code), and rural-urban commuting area codes were used to classify the place of residence into rural or urban areas. The respondents were also asked about their oldest care recipient’s age and place of residence (rural or urban areas classified based on the care recipient’s zip code). Respondents who reported being aware of their care recipient’s chronic conditions were also asked about their care recipients’ cognitive dysfunction (eg, dementia) and sensory impairment (eg, severe vision or hearing problems).

**Statistical Analysis**

Frequencies and percentages or means and SDs were used to describe the caregivers’ and their care recipients’ background information and use of technology devices and functions. Cohen κ statistics were estimated to examine the concordance of technology use among caregivers and care recipients. The magnitude of matching was classified into poor (κ<0.20), fair (κ=0.21-0.40), moderate (κ=0.41-0.60), good (κ=0.61-0.80), and very good (κ=0.81-1.00) matching categories. Along with Cohen κ coefficient estimation, the McNemar test was performed to compare the marginal proportions of caregivers and care recipients using or not using each technology. Next, separate multivariable Poisson regression analyses were performed to predict the total number of devices and functions used among caregivers based on caregivers’ age, sex, race and ethnicity, household income, education, and place of residence. Separate multivariable Poisson regression analyses were performed to predict the total number of devices and functions used among care recipients based on the care recipients’ age, place of residence, cognitive dysfunction, and sensory impairment and the total number of devices and functions used among caregivers. Only 438 (90.1%) out of 486 respondents were aware of their care recipients’ chronic conditions; therefore, the regression models for predicting care recipients’ technology use included a smaller sample size than the regression models for predicting caregivers’ technology use. All statistical analyses were performed using SAS 9.4 (SAS Institute Inc), and a significance level of .05 was used.

**Ethical Considerations**

The informed consent document was integrated at the outset of the web-based survey, and only those who agreed to participate proceeded to the subsequent sections of the survey. Given the web-based nature of the study, a waiver of documentation of informed consent was requested. Upon the completion of the study, any personally identifiable information (zip code) was deleted and age was truncated to 90 years old. Within Qualtrics, participant stipends were integrated into the survey, and each participant received a stipend upon the completion of the study. Based on the estimate provided Qualtrics, each participant was paid between US $7 and US $8. The study has been reviewed and approved by the Texas A&M University Institutional Review Board (IRB2019-1128M).

**Results**

**Study Participants**

The average age was 60.8 (SD 12.11) years for caregivers and 74.9 (SD 11.61) years for their oldest care recipient (Table 1). The majority of caregivers were female (363/485, 74.8%) and non-Hispanic White (331/483, 68.5%). Nearly 50% (241/486) had a household income less than US $50,000, and 20.6% (100/486) had high school or lower educational attainment. In all, 9.1% (44/483) of caregivers and 9.5% (46/482) of care recipients resided in rural areas. Of the 438 care recipients with available information, 43.2% (n=189) had cognitive dysfunction and 32.9% (n=144) had sensory impairment. On average, the caregivers used 3.4 devices and 4.2 functions, and their oldest care recipients used 1.8 devices and 1.6 functions.
Table. Characteristics of caregivers and care recipients and their technology use (n=486).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Caregivers</th>
<th>Care recipients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y), mean (SD)</td>
<td>60.8 (12.11)</td>
<td>74.9 (11.61)</td>
</tr>
<tr>
<td>Female, n/N (%)</td>
<td>363/485 (74.8)</td>
<td>N/Aa</td>
</tr>
<tr>
<td>Non-Hispanic White, n/N (%)</td>
<td>331/483 (68.5)</td>
<td>N/A</td>
</tr>
<tr>
<td>Household income less than US $50,000, n/N (%)</td>
<td>241/486 (49.6)</td>
<td>N/A</td>
</tr>
<tr>
<td>High school or lower educational attainment, n/N (%)</td>
<td>100/485 (20.6)</td>
<td>N/A</td>
</tr>
<tr>
<td>Rural residence, n/N (%)</td>
<td>44/483 (9.1)</td>
<td>46/482 (9.5)</td>
</tr>
<tr>
<td>Having cognitive dysfunction, n/N (%)b</td>
<td>N/A</td>
<td>189/438 (43.2)</td>
</tr>
<tr>
<td>Having sensory impairment, n/N (%)b</td>
<td>N/A</td>
<td>144/438 (32.9)</td>
</tr>
<tr>
<td>Number of devices used, mean (SD)c</td>
<td>3.4 (1.35)</td>
<td>1.8 (1.49)</td>
</tr>
<tr>
<td>Number of functions used, mean (SD)d</td>
<td>4.2 (1.73)</td>
<td>1.6 (1.92)</td>
</tr>
</tbody>
</table>

aN/A: not available.
bCognitive function and sensory impairment information was only available for care recipients and reported by 556 (88.8%) out of 626 total eligible caregivers.
cNumber of devices used ranged from 0 to 7.
dNumber of functions used ranged from 0 to 8.

Comparing Technology Use Among Caregivers and Care Recipients

McNemar tests showed that significantly greater proportions of caregivers used all examined technologies than their care recipients (all \( P < .05 \)), with the exception of the medication alerts or tracking function (\( P = .45 \); Table 2). \( \kappa \) coefficients ranged from 0.09 to 0.42 (Table 3), indicating a poor to moderate degree of matching (ie, concurrent use or no use) of technology among caregivers and care recipients. For example, there were 84% (404/481) of dyads in which the caregiver used a computer, yet there were only 31.2% (150/481) of dyads in which both the caregiver and care recipient used a computer.
Use of different devices and functions among caregivers and care recipients (n=486).

Table 1. Use of different devices and functions among caregivers and care recipients (n=486).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Caregivers, n/N (%)</th>
<th>Care recipients, n/N (%)</th>
<th>P valuea</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Devices</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell phone</td>
<td>268/481 (55.7)</td>
<td>231/481 (48)</td>
<td>.002</td>
</tr>
<tr>
<td>Smartphone</td>
<td>396/480 (82.5)</td>
<td>209/480 (43.5)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Tablet</td>
<td>263/481 (54.7)</td>
<td>112/481 (23.3)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Computer</td>
<td>404/481 (84)</td>
<td>160/481 (33.3)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>E-reader</td>
<td>98/480 (20.4)</td>
<td>44/480 (9.2)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Voice-activated assistant</td>
<td>143/481 (29.7)</td>
<td>74/481 (15.4)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Wearables for activity tracking</td>
<td>73/481 (15.2)</td>
<td>34/481 (7.1)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Functions

<table>
<thead>
<tr>
<th>Variables</th>
<th>Use by both caregivers and care recipients, n/N (%)</th>
<th>Nonuse by both caregivers and care recipients, n/N (%)</th>
<th>Cohen κ coefficienta</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Devices</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cell phone</td>
<td>180/481 (37.4)</td>
<td>162/481 (33.7)</td>
<td>0.42</td>
</tr>
<tr>
<td>Smartphone</td>
<td>194/480 (40.4)</td>
<td>69/480 (14.4)</td>
<td>0.17</td>
</tr>
<tr>
<td>Tablet</td>
<td>87/481 (18.1)</td>
<td>193/481 (40.1)</td>
<td>0.20</td>
</tr>
<tr>
<td>Computer</td>
<td>150/481 (31.2)</td>
<td>67/481 (13.9)</td>
<td>0.11</td>
</tr>
<tr>
<td>E-reader</td>
<td>26/480 (5.4)</td>
<td>364/480 (75.8)</td>
<td>0.27</td>
</tr>
<tr>
<td>Voice-activated assistant</td>
<td>58/481 (12.1)</td>
<td>322/481 (66.9)</td>
<td>0.42</td>
</tr>
<tr>
<td>Wearables for activity tracking</td>
<td>16/481 (3.3)</td>
<td>390/481 (81.1)</td>
<td>0.22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>Use by both caregivers and care recipients, n/N (%)</th>
<th>Nonuse by both caregivers and care recipients, n/N (%)</th>
<th>Cohen κ coefficienta</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Functions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td>168/481 (34.9)</td>
<td>64/481 (13.3)</td>
<td>0.12</td>
</tr>
<tr>
<td>Ride-sharing</td>
<td>20/480 (4.2)</td>
<td>376/480 (78.3)</td>
<td>0.25</td>
</tr>
<tr>
<td>Online shopping</td>
<td>127/481 (26.4)</td>
<td>82/481 (17)</td>
<td>0.12</td>
</tr>
<tr>
<td>Online banking</td>
<td>105/480 (21.9)</td>
<td>108/480 (22.5)</td>
<td>0.11</td>
</tr>
<tr>
<td>Navigation</td>
<td>75/479 (15.7)</td>
<td>132/479 (27.6)</td>
<td>0.09</td>
</tr>
<tr>
<td>Online entertainment</td>
<td>122/480 (25.4)</td>
<td>203/480 (42.3)</td>
<td>0.39</td>
</tr>
<tr>
<td>Medication alerts or tracking</td>
<td>22/481 (4.6)</td>
<td>397/481 (82.5)</td>
<td>0.34</td>
</tr>
<tr>
<td>Physical activity tracking</td>
<td>20/481 (4.2)</td>
<td>370/481 (76.9)</td>
<td>0.22</td>
</tr>
</tbody>
</table>

P value from the McNemar test examining whether there is a statistically significant difference in the proportions of caregivers and care recipients using or not using each technology.

Poor (κ<0.20), fair (κ=0.21-0.40), moderate (κ=0.41-0.60), good (κ=0.61-0.80), and very good (κ=0.81-1.00) matching.

https://aging.jmir.org/2024/1/e50759
**Correlates of Technology Use Among Caregivers**

The multivariable Poisson regression analyses suggested that the expected number of devices used among caregivers with household incomes less than US $50,000 was 0.88 times the number of devices used among those with household incomes of US $50,000 or more (\(P=.02\); Table 4). Using a separate regression analysis, the results indicated that the adjusted mean number of functions used among caregivers decreased by 0.8% for every 1-year increase in the caregivers’ age (\(P<.001\)). Additionally, the adjusted mean number of functions used among caregivers was negatively associated with household income (\(b=-0.097; \ P=.04\)) and educational attainment (\(b=-0.188; \ P=.002\)). The estimated number of functions used among caregivers in the lower household income and lower educational attainment categories was significantly lower than that of caregivers with higher socioeconomic status.

**Table**. Multivariable Poisson regression analysis for predicting the total number of devices and functions used among caregivers (n=486).

<table>
<thead>
<tr>
<th>Outcome and variables</th>
<th>b(^a) (SE)</th>
<th>IRR(^b) (95% CI)</th>
<th>(P) value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of devices</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.001 (0.002)</td>
<td>1.001 (0.997-1.005)</td>
<td>.67</td>
</tr>
<tr>
<td>Female</td>
<td>0.035 (0.059)</td>
<td>1.036 (0.923-1.162)</td>
<td>.55</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>-0.026 (0.054)</td>
<td>0.975 (0.877-1.083)</td>
<td>.63</td>
</tr>
<tr>
<td>Household income less than US $50,000</td>
<td>-0.130 (0.053)</td>
<td>0.878 (0.791-0.975)</td>
<td>.02</td>
</tr>
<tr>
<td>High school or lower educational attainment</td>
<td>-0.086 (0.066)</td>
<td>0.918 (0.806-1.045)</td>
<td>.20</td>
</tr>
<tr>
<td>Rural residence</td>
<td>-0.079 (0.094)</td>
<td>0.924 (0.769-1.110)</td>
<td>.40</td>
</tr>
<tr>
<td><strong>Number of functions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.008 (0.002)</td>
<td>0.992 (0.989-0.996)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Female</td>
<td>-0.017 (0.053)</td>
<td>0.983 (0.886-1.090)</td>
<td>.74</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>-0.046 (0.048)</td>
<td>0.955 (0.869-1.049)</td>
<td>.34</td>
</tr>
<tr>
<td>Household income less than US $50,000</td>
<td>-0.097 (0.048)</td>
<td>0.908 (0.826-0.998)</td>
<td>.04</td>
</tr>
<tr>
<td>High school or lower educational attainment</td>
<td>-0.188 (0.061)</td>
<td>0.829 (0.735-0.934)</td>
<td>.002</td>
</tr>
<tr>
<td>Rural residence</td>
<td>-0.068 (0.085)</td>
<td>0.934 (0.792-1.103)</td>
<td>.42</td>
</tr>
</tbody>
</table>

\(^a\)b: regression coefficient.

\(^b\)IRR: incidence rate ratio.

**Correlates of Technology Use Among Care Recipients**

For care recipients, age and cognitive dysfunction were negatively associated with number of devices (\(b=-0.024; \ P<.001\) and \(b=-0.394; \ P<.001\), respectively) and functions (\(b=-0.032; \ P<.001\) and \(b=-0.370; \ P=.002\), respectively; Table 5). In contrast, caregivers’ technology use was positively associated with care recipients’ technology use (\(b=0.184; \ P<.001\) for devices and \(b=0.238; \ P<.001\) for functions; Table 5).
Table. Multivariable Poisson regression analysis for predicting the total number of devices and functions used among care recipients (n=438).

<table>
<thead>
<tr>
<th>Outcome and variables</th>
<th>b (SE)</th>
<th>IRR (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Number of devices</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.024 (0.003)</td>
<td>0.976 (0.970-0.983)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Rural residence</td>
<td>-0.102 (0.125)</td>
<td>0.903 (0.707-1.154)</td>
<td>.41</td>
</tr>
<tr>
<td>Cognitive dysfunction</td>
<td>-0.394 (0.080)</td>
<td>0.675 (0.577-0.789)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Sensory impairment</td>
<td>0.002 (0.082)</td>
<td>1.002 (0.854-1.176)</td>
<td>.98</td>
</tr>
<tr>
<td>Number of devices used by caregiver</td>
<td>0.184 (0.027)</td>
<td>1.202 (1.142-1.267)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Number of functions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-0.032 (0.005)</td>
<td>0.969 (0.960-0.979)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Rural residence</td>
<td>-0.069 (0.186)</td>
<td>0.934 (0.648-1.344)</td>
<td>.71</td>
</tr>
<tr>
<td>Cognitive dysfunction</td>
<td>-0.370 (0.121)</td>
<td>0.691 (0.545-0.876)</td>
<td>.002</td>
</tr>
<tr>
<td>Sensory impairment</td>
<td>-0.038 (0.123)</td>
<td>0.963 (0.756-1.225)</td>
<td>.76</td>
</tr>
<tr>
<td>Number of devices used by caregiver</td>
<td>0.238 (0.053)</td>
<td>1.269 (1.189-1.354)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

a,b: regression coefficient. 

bIRR: incidence rate ratio.

Discussion

Principal Findings

Based on the data analysis and results analyzed in the above section, some key findings provide a fuller and more specific understanding of older adults’ and their caregivers’ use of digital technologies and highlight the contextual factors that may either affect resistance or enhance accessibility and receptivity of technology-based interventions in a broader caregiver population.

Key findings and reflections on current and parallel research are presented below.

Use of Technology and Comparison

This study examined unpaid caregivers’ use of various technology and their reports of their older adult care recipient’s use. It was observed that a greater proportion of caregivers reported using technologies themselves than being used by their care recipients. In our study, although 82.5% (396/480) and 84% (404/481) of caregivers reported using smartphones and computers, only 43.5% (209/480) and 33.3% (160/481) of older adults used smartphones and computers, respectively. This implies greater access to technology by caregivers than their older adult care recipients. Furthermore, although caregivers reported a higher number of functions used than the number of device types used (eg, 3.4 devices and 4.2 functions), their older care recipients reported a fewer number of functions than the number of device types used (eg, 1.8 devices and 1.6 functions). These findings imply that caregivers are more likely to maximize the potential of a technology than their care recipients, who are typically older and in poorer health [20,21]. Our findings support the 2020 AARP tech trend report [5], which pointed out that despite older adults’ high engagement with their devices, “many are not using the technology to its full potential.” For example, fewer than half of smart home assistant or smart speaker owners used the device daily according to the 2019 national survey [5]. Although the proportion increased to about 57% in 2023 [35], the rate has still remained relatively low. The value of technology in enabling aging in place and reducing caregiver burden will only become further enhanced when these devices can be used to their fullest potential by older adults.

The age-related discrepancy in technology adoption is likely to be associated with skills in using technology but also with attitudes related to technology. For example, “perceived needs” is an important attitudinal factor in behavior adoption [36]. In this study, there was the lack of statistically significant differences between caregivers and their care recipients in medication management technology. These findings are consistent with research by both Abrashkin et al [28] and Portz et al [37], who also found that older adults in an advanced illness management program had significantly less access to and confidence in using technology (eg, computer, internet, tablet, and cell phone) than the program enrollees’ caregivers, except for medical alert devices such as medication management systems [28,37]. While the reasons underlying these findings have not been fully examined yet, this could be related to a similar level of perceived needs by both parties.

Correlates of Technology Use in Caregivers

Among caregivers, older age and lower socioeconomic factors (household income and education) were negatively associated with the number of technological devices or functions used. This finding is consistent with extant literature [38-41]. Our study further explores this relationship by revealing a difference between devices and functions. Although the number of technological devices used by caregivers was only significantly associated with household income, the number of technological functions in use was associated with multiple factors (ie, age, household income, and education).
Regarding ownership or access to technological devices, the age and socioeconomic aspects of the digital divide seem to be narrowing [5]. However, the age and socioeconomic aspects of the digital divide appear to remain in technology use [42]. Among the more critical issues facing many older adults in using digital devices such as smartphones and tablets, they lack the basic digital literacy required to use multimedia interactive devices with touchscreen technology [43,44], and their digital literacy level is likely to diminish with age [35]. Assumptions in the past were that if access to devices and basic training were provided, the “grey span” of the digital divide could be eliminated. For older adults, each new operating system revision or interface for existing devices can be a traumatic event, as what worked before no longer does. Cao et al [45] described the information overload and system feature overload of new digital applications that resulted in increased fatigue and technostress of the older adult users, further increasing their resistance to technology adoption. These findings align with the AARP national survey findings indicating a limited set of tasks performed by older adults on technology [5,35]. This continuing of the digital divide can disproportionately impact caregiving for older care recipients, especially those whose primary caregivers are often older adults.

**Correlates of Technology Use in Care Recipients**

Our study found that, for care recipients, age and cognitive dysfunction were negatively associated with the number of devices and functions. In contrast, care recipients’ use of a device or function was positively associated with the use of the technology among their caregivers. Along with the previous findings about the correlates of technology use in caregivers, these findings align with Baishya and Samalia’s [46] assertion that technology adoption is contextual. This study’s findings emphasize the need for additional research to identify and understand the contextual factors to enhance the accessibility and receptivity of technology-based interventions in a broader aging community.

**Limitations**

This study is not without limitations. Efforts were made to diversify the convenience Qualtrics panel sample by specifying the proportion of sample characteristics of respondents to ensure heterogeneity in key factors such as age, gender, race and ethnicity, and geographic region. However, this study did not use probability sampling and is subject to a potential nonprobability sampling bias. For example, given the nature of web-based recruitment, the study population already has access to the internet. Although the majority of caregivers have access to the internet (eg, 78% in 2018 [47]), the study findings may not be generalizable to a group of caregivers without access to the internet. Additionally, this study used proxy responses for care recipients’ technology use (ie, as reported by their caregivers). Further, some key sociodemographic characteristics of care recipients were not collected.

Our restriction to unpaid caregivers who provided at least 8 weekly hours of care was intended to ensure that caregivers were familiar with their care recipients. However, it is important to note that previous literature indicates that caregivers tend to underestimate care recipients’ physical and cognitive functions and certain activities [48,49] and may also be imprecise in reporting technology use. In this study, many care recipients had cognitive impairments or sensory impairments, which might have adversely affected caregivers’ perception of care recipients’ user capacity and actual use of technology. However, we also note that many older persons do experience these functional limitations and impairments and their inclusion is important for reflecting health conditions and technology use in this population, albeit recognizing reporting limitations. This study was not able to differentiate the cross-use or whether the technology was used explicitly for caregiving functions. Future studies can benefit from the specification of the purpose of different technologies related to caregiving tasks and more precise measurement of care recipients’ technology use, such as daily diary use of technology devices and functional use over a specified period of time or digital tracking of technology use.

This study was conducted prior to the onset of the COVID-19 pandemic; hence, it does not reflect what might have changed in attitude toward or use of technologies, as well as the evolution of technologies. However, it differentiates between technology use and function among both caregivers and care recipients and provides important insights related to disparities in access to technology, which was a critical factor in access to health care and other social services during the COVID-19 pandemic. Furthermore, despite the increase in technology use in older adults since the COVID-19 pandemic, this study’s findings align with the more recent report on older adults’ technology use pattern regarding the use of functions and socioeconomic correlates (eg, income) [35].

**Conclusions**

With the increasing use of technology solutions for caregiving that are becoming available on the market, it is important to be aware of factors associated with the current digital divide in technology use—both in terms of the number and diversity of devices and their functional use. It is critical to look forward to what the future might hold regarding the technology being used to reduce caregiver burden and enhance care recipients’ health, independence, and quality of life. A digital divide among older adults can exacerbate greater health disparity since technology is a powerful source for obtaining information and communicating with health care and social service providers [50].

A major finding from this study was the existence of significant disparities in the use of technological devices and functions among caregivers and their older adult care recipients. Among caregivers, significant differences were observed in technology use based on age and socioeconomic factors. In addition, this study suggested that caregivers’ technology use is an enabling factor for older care recipients’ technology use, independent of advanced age and cognitive impairment, which depressed use, indicating pathways for clinical intervention.

This study was conducted before the onset of the COVID-19 pandemic, which demonstrated the growing importance of connecting on the web for basic health care. The extent to which observed relationships between individual and socioeconomic moderators and technology use have changed since COVID-19 is a question for further study. The importance of technology...
use has become more salient during the recommended “physically distant stay at home” orders for older adults to stay socially connected with loved ones or professional social connectors, whether living at home, in assisted living facilities, or even nursing homes [51]. Additionally, familiarity with or access to technology can facilitate or act as a barrier to obtaining COVID-19 vaccinations. For example, the reach of a digital platform to track vaccinations and make follow-up interactions among older adult populations, who would benefit greatly from such technology, will depend largely upon the extent telecommunication or telehealth is used or accepted by older adults or their caregivers.

In summary, this study adds to the rapidly expanding field of technology in the health and aging realm by describing potential contextual factors in technology use, which may contribute to the disparities in technology use among older adults and their care recipients [52,53]. Further efforts are needed to expand the understanding of how these contextual factors contribute to technology adoption among caregivers and their care recipients and the benefits and costs of such technological innovations [54]. Especially relevant is how social workers, health professionals, educators, and the community can facilitate and maintain appropriate use of new and emerging technology for critical interactions normally and enable access to the needed caregiver and social resources during the COVID-19 pandemic or after it subsides. Furthermore, future research could gain additional benefits by concentrating on broader categories of functions. This approach would enable a more targeted investigation into particular functions related to specific outcomes, such as economic functions and financial health.

Acknowledgments

We thank all the caregivers who participated in our web-based survey, which formed the basis of this research. The caregiver survey was funded by contributions from DVD Associates LLC, Clairvoyant Networks Inc, and The Texas A&M Center for Population Health and Aging. Clairvoyant Networks Inc was not involved in the manuscript’s conceptualization, data analyses, interpretation, or development.

Authors’ Contributions

MGO lead the study conceptualization and manuscript development. SL conducted all statistical analyses. All authors contributed to the study conceptualization, data collection, and development and revision of the manuscript.

Conflicts of Interest

None declared.

References


https://aging.jmir.org/2024/1/e50759
11. Abrahms S. New technology could allow you or your parents to age at home. AARP Bulletin. 2014 Mar 19. URL: [https://www.aarp.org/home-family/personal-technology/info-2014/is-this-the-end-of-the-nursing-home.html](https://www.aarp.org/home-family/personal-technology/info-2014/is-this-the-end-of-the-nursing-home.html) [accessed 2024-03-08]


Assessing the Impact of Internet Skills on Depressive Symptoms Among Chinese Middle-Aged and Older Adults: Cross-Sectional Instrumental Variables Analysis

Aruhan Mu, MA; Zhiyong Liu, PhD
School of Medicine and Health Management, Huazhong University of Science and Technology, Wuhan, China

Corresponding Author:
Zhiyong Liu, PhD

Abstract

Background: The potential benefits of IT for the well-being of older adults have been widely anticipated. However, findings regarding the impact of internet use on depressive symptoms are inconsistent. As a result of IT’s exponential growth, internet skills have supplanted internet access as the source of the digital divide.

Objective: This study evaluates the effect of internet skills on depressive symptoms through an instrumental variables (IV) approach.

Methods: Data from the China Health and Retirement Longitudinal Study’s wave 4 (2018) were used. This included 16,949 community residents aged 45 years and older. To overcome the endogeneity issue, we used an IV approach.

Results: Our results reveal the emergence of a second-level digital divide, the disparity in internet skills, among Chinese middle-aged and older adults. Linear regression suggests that a 1% increase in internet skills is associated with a 0.037% decrease in depressive symptoms ($\beta = -0.037, SE = 0.009$), which underestimates the causal effect. As expected, internet skills are an endogenous variable ($F$ test $P$ value < .001). IV regressions indicate that a 1% increase in internet skills reduces 1.135% (SE 0.471) to 1.741% (SE 0.297) of depressive symptoms. These 2 IV are neither weak ($F_{-1} = 16.7$ and 28.5; both >10) nor endogenous (Wu-Hausman test $P$ value of .10; >.05 or >.01).

Conclusions: Better mental health is predicted through improved and higher internet skills. Consequently, residents and policy makers in China should focus on bridging the digital divide in internet skills among middle-aged and older adults.

(JMIR Aging 2024;7:e50880) doi:10.2196/50880

KEYWORDS
internet skills; depression; second-level digital divide; instrumental variables

Introduction

Background

By 2050, there will be 2.1 billion individuals aged 60 years or older, with 80% of them living in low- and middle-income countries [1]. Late-life depression is a major public health challenge for this population because of its high prevalence and poor outcomes [2,3]. China, one of the low- and middle-income countries, is predicted to reach nearly 400 million older adults by 2050 [4]. Approximately 20%-31% of individuals aged 45 years or older in China had depression in 2015, with the highest risk among all age groups [5,6]. In 2013, the annual cost for individuals with mental disorders in China was US $3665, with depression and depressive symptoms accounting for 54% of the total [7,8]. With a rapidly aging population, late-life depression may damage the well-being of older adults; it also brings more burden to the family as well as the community. China accounts for nearly 17% of the global mental health burden [9]. In contrast, China’s mental health system has only 8.75 mental health workers per 100,000 residents [10], which calls for novel approaches with low cost and wide access to reduce depressive symptoms among middle-aged and older adults.

Given that the internet can improve one’s mental health by reducing social isolation and loneliness, it has attracted marked attention among researchers [11]. Prior work has investigated the correlation between depression symptoms and the frequency, type, and purpose of internet use, and identified potential mechanisms from the perspective of social connectedness [12-16]. Overall, these studies suggest that internet use among older adults offers new opportunities to prevent, support, and treat late-life depression in family and community contexts.

We aim to complement this research topic in 2 ways. First, previous studies have mainly measured internet use from the perspective of the first-level digital divide, that is, whether respondents have access to the internet [12,16,17]. With the rapid penetration of IT, internet skills (the capability of internet use) have replaced internet access as the most essential variable in characterizing IT use [18-20]. We believe that we have now
entered the second level of the digital divide and thus chose internet skills to measure internet use. Second, the existing literature has derived results mainly from correlation analysis and yielded mixed evidence [13,16,21,22]. However, the findings of correlation analysis cannot provide interventional insights; only causal findings can do so [23]. Several modern economic approaches have been used in empirical studies to estimate the causal relationship, such as the instrumental variables (IV) approach. Joshua D Angrist, who shares the 2021 Nobel Prize for methodological contributions to the analysis of causal relationships, adopted and developed the IV approach to quantify the impact of educational attainment on wage growth [24]. In the face of rising public health challenges in an aging society, policy makers and the public need robust insights. Therefore, our goal is to uncover the causal effects of internet skills on depressive symptoms in middle-aged and older adults and provide practical implications for promoting mental health through daily IT use.

Overall, IT-based mental health promotion programs could be an essential means to address the high prevalence of depression in older adults when they lack specialized medical resources [25]. As the internet spreads to older adults, we should pay more attention to differences in internet skills rather than in internet access [20,26]. We should also focus on causal effects to provide more valuable evidence for public health practices [23]. Therefore, this study examines the causal effect of internet skills on depressive symptoms among middle-aged and older Chinese adults by using data from the 2018 China Health and Retirement Longitudinal Study (CHARLS). We used the IV approach to address the endogeneity issue and guide daily internet use practices among middle-aged and older adults to improve their mental health.

Literature Review

Internet Use and Depressive Symptoms Among Middle-Aged and Older Adults

Numerous studies have examined the relationship between internet use and depressive symptoms in older adults using national representative data sets. Table S1 in Multimedia Appendix 1 summarizes the key findings from these studies. Several insights are worth noting. First, analyses of the correlation between internet use and depressive symptoms yielded inconsistent conclusions. Jun and Kim [27] reported in 2015 that internet use was associated with lower levels of depressive symptoms among middle-aged and older Chinese adults by using data from the 2018 China Health and Retirement Longitudinal Study (CHARLS). We used the IV approach to address the endogeneity issue and guide daily internet use practices among middle-aged and older adults to improve their mental health.

Internet Skills

The digital divide refers to certain groups having better opportunities than others to benefit from IT. Prior studies have suggested three stages of the digital divide [18,20,26]: (1) economic divide, which implies that some people cannot afford access to IT; (2) usability divide, which emphasizes that IT remains so complicated that some people cannot use it even if they can afford it; and (3) empowerment divide, which refers to inequality of outcomes after IT use. As internet use becomes prevalent among older adults, it is increasingly important to look at who uses the internet and distinguish their internet skills [18,33]. For depression among middle-aged and older adults, the existing literature has primarily discussed the first-level digital divide, which refers to the impact of internet use on depressive symptoms. We further examine the second-level digital divide, namely how internet skills influence depressive symptoms among middle-aged and older adults.

IV of Internet Skills

On examining how internet skills affect depressive symptoms, endogenous issues may be raised for two main reasons: (1) internet skills and depressive symptoms are simultaneously affected by unobserved factors such as personality traits, and (2) mentally healthier individuals are more prone to use the internet and have better internet skills. We have used the IV approach to handle potential endogeneity and obtain reliable causal effects.
A few studies have applied the IV approach to capture the effect of internet use. Hong and Chang [34] used the geographical distance to the nearest telecommunication station for each household as an instrumental variable to estimate the impact of internet use on household income at forestry farms in Fujian Province in China. Gao et al [35] applied provincial internet penetration rates as an instrumental variable to capture the effect of computer penetration on Chinese rural farmers’ income. Nie et al [31] used the number of provincial internet broadband access terminals as an instrumental variable to examine the relationship between internet use and depressive symptoms among 16- to 60-year-old Chinese individuals.

This study adopts the above-described “resources accessibility” approach by using the following 2 IV: mobile phone penetration and the performance of government websites. Specifically, higher mobile phone penetration in a city implies that residents have easier access to IT resources, which implies a higher level of IT acceptance and internet skills [31,35]. Similarly, better performance of government website operations indicates higher informatization of the city, which implies higher internet skills of residents in that city [34]. Furthermore, there is no evidence that city-level mobile phone penetration or government website performance is directly associated with individual depressive symptoms. Overall, mobile phone penetration and government website performance correlate with each resident’s internet acceptance and skills, while not being directly correlated with individual depressive symptoms. Therefore, these variables satisfy the principle of IV selection logically [24,36]. Furthermore, when considering the population of older adults in each city, both variables are multiplied by the proportion of the city’s population aged 60 years and older.

### Methods

#### Sample and Data Collection

We analyzed data from CHARLS, a nationwide survey designed to provide comprehensive and high-quality data on the demographics, household characteristics, health status and functioning, work, and retirement information of Chinese residents aged 45 years and older [37]. CHARLS is a longitudinal study that used a 4-stage, stratified, cluster sampling design to enroll community-dwelling residents from 450 villages and 150 counties in 28 provinces in China. The national baseline study (wave 1) was conducted in 2011. The last public survey (wave 4) was conducted in 2018, with information obtained from 19,816 respondents.

Since IV regression can investigate the causal effects of independent variables on dependent variables in cross-sectional data, we selected 16,949 participants from CHARLS wave 4. The sampling process is as follows: (1) respondents aged 45 years and older (excluded 178 samples); (2) provided information on internet skills and depression (excluded 2689 samples).


#### Variables

**Depression**

The 10-item Center for Epidemiologic Studies Depression Scale (CES-D-10) is used to examine depressive symptoms. The respondents were asked about their positive feelings, negative emotions, and somatic symptoms during the past week. Scores for each question ranged from 0 to 30, with high scores indicating severe depressive symptoms. Our study considers 12 as a cutoff point to describe the prevalence of depression [38], using the CES-D-10 score in IV regression.

**Internet Skills**

New questions about internet skills have been added in wave 4 of CHARLS. Respondents were asked whether they would use the following web-based functions on their mobile phones: (1) chat on social media (such as WeChat); (2) post on social media (such as WeChat moments); and (3) mobile payments (such as Alipay or WeChat). Respondents provided binary responses, denoting “yes” or “no.” Subsequently, based on the perceived complexity associated with these distinct functions, this study assigned numerical scores to gauge respondents’ internet skills. Respondents capable of using mobile payment systems were assigned a score of 9, those adept at posting on social media received a score of 5, individuals proficient in social media chat were assigned a score of 1, while those abstaining from all the aforementioned functionalities received a score of 0. The cumulative internet skills score was derived by evaluating respondents’ competencies across various functions, within a scale spanning from 0 to 15 points. Higher scores herein signify enhanced internet aptitude.

To reinforce the dependability of our findings, we incorporated an additional measurement strategy, using both a Likert scale with scores ranging from 1 to 5 and a cumulative scoring method as alternative approaches for assessing internet skills. An IV analysis was conducted using these supplementary metrics of internet skills. The inferences drawn from this analysis are in concordance with the foundational model’s findings. Detailed information on the measurements and the corresponding results can be found in Tables S1-S4 in Multimedia Appendix 1.

#### Mobile Phone Penetration and Performance of Government Websites

The first instrumental variable is mobile phone penetration: the number of mobile phone subscribers at the end of the year in a city multiplied by the proportion of individuals aged 60 years and older in each city. Another instrumental variable is the performance of government websites: the score is the sum of the operational scores of each city’s government website, ranging from 50 to 100.

#### Potential Confounding Variables

Potential confounding variables are shown in Table 1.
Table. Definition/codes of the potential confounding variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Codes/definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>Continuous variable</td>
</tr>
<tr>
<td>Gender</td>
<td>0=male; 1=female</td>
</tr>
<tr>
<td>Marital status</td>
<td>0=single (divorced, widowed, or single); 1=partnered (married or partnered)</td>
</tr>
<tr>
<td>Retirement</td>
<td>0=no; 1=yes</td>
</tr>
<tr>
<td>Education</td>
<td>0=less than lower secondary (no formal education, did not finish primary school but can read, private tutoring, elementary school, or middle school); 1=upper secondary (high school or vocational school); 2=tertiary (2- or 3-year college, college graduate, or postgraduate degree)</td>
</tr>
<tr>
<td>Total household per capita consumption</td>
<td>Total household consumption or number of people living in the household (skewed distribution, logarithmically transformed)</td>
</tr>
<tr>
<td>Residency</td>
<td>0=urban; 1=rural</td>
</tr>
<tr>
<td>Ever had a memory problem</td>
<td>Ever had doctor-diagnosed memory-related diseases, including dementia, brain atrophy, and Parkinson disease (0=no; 1=yes)</td>
</tr>
<tr>
<td>Ever had a psychological problem</td>
<td>Ever had doctor-diagnosed psychiatric problems, such as emotional, nervous, or psychiatric problems (0=no; 1=yes)</td>
</tr>
<tr>
<td>Mobility</td>
<td>A 9-item summary of any difficulty with mobility activities. The mobility activities are walking 100 m, climbing several flights of stairs, getting up from a chair, stooping or kneeling or crouching, extending arms up, lifting 5 kg, and picking up a small coin. Continuous variable: 0-27 (skewed distribution, logarithmically transformed)</td>
</tr>
</tbody>
</table>

Ethical Considerations

This investigation constitutes a secondary analysis of publicly available data sets; hence, prior registration was not deemed necessary. The foundational data used herein are obtainable via the official portal of the CFPS [39]. Ethical approval for the CFPS was granted from the Biomedical Ethics Committee at Peking University (IRB00001052-14010).

Statistical Analysis

To examine the effect of internet skills on depressive symptoms, we performed an IV regression to control for possible endogeneity issues. Given that both the dependent and independent variables have skewed distributions, it is meaningful to understand the impact of internet skills on depressive symptoms in terms of percentage change. Thus, we estimate a log-log specification presented in the two-stage least squares model as follows:

\[
\begin{align*}
\text{(1)} & \quad \text{ces-d-10} = \beta_0 + \beta_1 \text{internet skills} + \beta_2 \text{potential confounding variables} + \beta_3 \text{city mobile phone penetration} + \epsilon_n \\
\text{(2)} & \quad \text{city mobile phone penetration} = \alpha_0 + \alpha_1 \text{govern government website performance} + \epsilon_m
\end{align*}
\]

where \( \epsilon_n \) is the error term.

We carried out regression analysis to evaluate the association of individual depressive symptoms with individuals’ internet skills and the mobile phone penetration in the city. For another instrumental variable, the first stage of the two-stage least squares model also can be written as follows:

\[
\text{(3)} & \quad \text{govern government website performance} = \gamma_0 + \gamma_1 \text{potential confounding variables} + \gamma_2 \text{city mobile phone penetration} + \epsilon_g
\]

Diagnostic tests including correlation, overidentification, and weak IV tests supported the validity of the aforementioned IV. The data were analyzed using R package AER (version 1.2; R Foundation for Statistical Computing).

Results

Descriptive Statistics

The demographic characteristics of the sample are shown in Table 2. There are 16,949 participants, and the mean age in wave 4 (2018) was 62.3 (SD 9.9) years. Most participants are female (n=8735, 51.5%), have a partner (n=14,672, 86.6%), have less than a lower secondary level of education (n=14,619, 86.2%), are currently employed (n=11,086, 66.5%), and are rural residents (n=10,095, 60.7%). In addition, internet skills have significant differences across demographics, including age, gender, and education levels.
Table. Characteristics of the selected respondents. The total percentage may not equal 100 due to rounding.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Full sample (n=16,949)</th>
<th>Scores for IT skills</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>0 (n=14,601, 86.1%)</td>
<td>1 (n=312, 1.8%)</td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>62.3 (9.9)</td>
<td>63.3 (9.8)</td>
<td>57.4 (7.7)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>8214 (48.5)</td>
<td>6944 (47.6)</td>
<td>138 (44.2)</td>
</tr>
<tr>
<td>Female</td>
<td>8735 (51.5)</td>
<td>7657 (52.4)</td>
<td>174 (55.8)</td>
</tr>
<tr>
<td>Marital status, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>2277 (13.4)</td>
<td>2114 (14.5)</td>
<td>23 (7.4)</td>
</tr>
<tr>
<td>Partnered</td>
<td>14,672 (86.6)</td>
<td>12,487 (85.5)</td>
<td>289 (92.6)</td>
</tr>
<tr>
<td>Education, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than lower secondary</td>
<td>14,619 (86.2)</td>
<td>13,204 (90.4)</td>
<td>241 (77.2)</td>
</tr>
<tr>
<td>Upper secondary</td>
<td>1939 (11.4)</td>
<td>1242 (8.5)</td>
<td>59 (18.9)</td>
</tr>
<tr>
<td>Tertiary</td>
<td>391 (2.3)</td>
<td>155 (1.1)</td>
<td>12 (3.8)</td>
</tr>
<tr>
<td>Total household per capita consumption&lt;sup&gt;c&lt;/sup&gt;, median</td>
<td>6240.8</td>
<td>5657</td>
<td>7400.2</td>
</tr>
<tr>
<td>Retirement status&lt;sup&gt;d&lt;/sup&gt;, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not retired</td>
<td>11,086 (66.5)</td>
<td>9527 (66.1)</td>
<td>208 (68.2)</td>
</tr>
<tr>
<td>Retired</td>
<td>5575 (33.5)</td>
<td>4877 (33.9)</td>
<td>97 (31.8)</td>
</tr>
<tr>
<td>Residential area&lt;sup&gt;e&lt;/sup&gt;, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>6530 (39.3)</td>
<td>5149 (35.8)</td>
<td>146 (48.2)</td>
</tr>
<tr>
<td>Rural</td>
<td>10,095 (60.7)</td>
<td>9231 (64.2)</td>
<td>157 (51.8)</td>
</tr>
<tr>
<td>Ever had memory problems&lt;sup&gt;f&lt;/sup&gt;, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>16,283 (97.9)</td>
<td>13,977 (97.8)</td>
<td>305 (99.0)</td>
</tr>
<tr>
<td>Yes</td>
<td>346 (2.1)</td>
<td>320 (2.2)</td>
<td>3 (1.0)</td>
</tr>
<tr>
<td>Ever had psychological problems&lt;sup&gt;g&lt;/sup&gt;, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>16,512 (98.9)</td>
<td>14,197 (98.9)</td>
<td>309 (99.7)</td>
</tr>
<tr>
<td>Yes</td>
<td>179 (1.1)</td>
<td>164 (1.1)</td>
<td>1 (0.3)</td>
</tr>
<tr>
<td>Mobility&lt;sup&gt;b&lt;/sup&gt;, mean (SD)</td>
<td>4.1 (5.1)</td>
<td>4.4 (5.3)</td>
<td>2.7 (3.7)</td>
</tr>
<tr>
<td>CES-D-10&lt;sup&gt;h&lt;/sup&gt; score, mean (SD)</td>
<td>8.4 (6.5)</td>
<td>8.8 (6.6)</td>
<td>7.2 (6.0)</td>
</tr>
<tr>
<td>Mobile phone user rate, mean (SD)</td>
<td>127.7 (127)</td>
<td>125.6 (125)</td>
<td>131 (133.6)</td>
</tr>
<tr>
<td>Government website performance, mean (SD)</td>
<td>63.8 (15.1)</td>
<td>63.5 (14.9)</td>
<td>64.2 (15.4)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Outcome of the Kruskal-Wallis test.
<sup>b</sup>Outcomes of the chi-square test.
<sup>c</sup>Missing data: n=2459.
<sup>d</sup>Missing data: n=288.
<sup>e</sup>Missing data: n=324.
<sup>f</sup>Missing data: n=320.
<sup>g</sup>Missing data: n=258.
<sup>h</sup>Missing data: n=44.
Table 1 reveals a distinct profile among individuals with elevated internet skill scores. These individuals are predominantly younger, hail from higher-income households (higher total household per capita consumption), are more likely to be male, have partners, are nonretired, and are predominantly urban residents. Additionally, this cohort is characterized by superior health outcomes, evident from lower incidences of memory-related issues, psychological problems, and reduced mobility challenges.

Further analysis of internet skill scores underscores a stark digital divide. A significant 86.1% of participants do not engage with the internet, thereby receiving a score of 0. Conversely, within the subset of internet users, those with advanced internet skills form the majority, representing 7.1% of the overall sample and 51.4% of internet users (1206 out of 2348 individuals). This disparity highlights not only a first-level digital divide but also a pronounced second-level digital divide within the middle-aged and older demographic, marked by a dual extremity of complete absence or high proficiency in internet skills.

Prevalence of Depression in Residents

The prevalence of depression varied among participants with different levels of internet skills. Respondents with lower-level internet skills are at higher risk of depression than those with higher-level internet skills (30.23% vs 20.51% vs 17.38% vs 13.24% vs 15.17%).

The Effect of Internet Skills on Depressive Symptoms in Residents

The association between internet skills and depressive symptoms was estimated using ordinary least squares (OLS) regression, and the effect of internet skills on depressive symptoms was subsequently estimated using IV regression (see Table 3). OLS regression suggests that a 1% increase in internet skills is associated with a 0.037% decrease in depressive symptoms (β=−.037, SE 0.009), which underestimates the causal effect. As expected, internet skills are an endogenous variable (F test P value <.001). IV regressions indicate that a 1% increase in internet skills leads to a 1.135% (SE 0.471) to 1.741% (SE 0.297) reduction in depressive symptoms. The 2 IV were neither weak (F test P value of .10; being >.05 or .01) nor endogenous (Wu-Hausman test P value of .10; being >.05 or .01). Detailed results of the controls can be found in Table S5 in Multimedia Appendix 1.
Sensitivity Analysis
We applied multiple imputation techniques to missing values, assuming that missing variables are missing at random. To examine the robustness of the results, we conducted OLS and IV regressions with missing values and compared the results. Our findings yielded no difference between complete cases and the prior sample. Detailed results can be found in Table S6 in Multimedia Appendix 1.

Discussion
This study explored the impact of internet skills on depressive symptoms. Our findings revealed that (1) internet skills are relatively low among Chinese middle-aged and older adults and (2) improvement of internet skills can reduce depressive symptoms.

Impact of Internet Skills on Depressive Symptoms
Our results suggest that for middle-aged and older Chinese adults, a 1% increase in their internet skills leads to a 1.1% to 1.7% reduction in depressive symptoms. Our main finding is consistent with those of Cotten et al [12,30], who reported a positive outcome of internet use reducing risk of depression by 20% to 33% among retired older adults in the United States. We used internet skills (0–15 points) to measure IT use rather than using dummy variables. In addition, since the CES-D-10 scores are a continuous variable, we could estimate the effect of internet skills on depressive symptoms.

Some studies have reported that internet use increases depressive symptoms in the context of China [31,32]. Nie et al [31] used the IV approach and 2010 CFPS data and found that internet use was associated with higher levels of depressive symptoms among Chinese individuals aged 16–60 years. Notably, in 2010, internet access among middle-aged and older Chinese individuals was limited [40], and some residents were even prejudiced and resistant to the internet [31]. These reasons may lead to a negative impact of internet use on depression. The existing literature predominantly addresses the first-level digital divide, examining the effect of internet use on depressive symptoms [31,32]. Our study expands on this by delving into the second-level digital divide, revealing that enhanced internet skills significantly mitigate depressive symptoms. This observation, divergent from prior studies, suggests a possible shift in the perception of the internet among the middle-aged and older Chinese individuals as it becomes more ingrained in various facets of society [31]. Proficiency in internet use potentially enables more effective usage [13,16,19], leading to positive psychosocial outcomes. For instance, improved internet skills facilitate web-based social interactions and content sharing, thus increasing social support and reducing isolation and loneliness [11,15,22,27]. Further skill enhancement allows these individuals to more efficiently perform daily activities such as shopping and bill payments on the web, fostering a greater sense of inclusion and autonomy in the digital era [14,16,19]. Such psychosocial benefits potentially alleviate depressive symptoms and bolster mental health.

Our study additionally revealed that populations with a higher SES and superior health tend to exhibit more advanced internet skills. While it has been acknowledged that SES can create disparities in resource access and health outcomes [29,33], our research indicates that the improvement of internet skills might mitigate depressive symptoms, even in scenarios where SES and health factors are consistent. This finding suggests that diminishing the second-level digital divide could play a role in lessening health disparities. Consequently, future research should delve into the potential mechanisms by which internet skills influence depressive symptoms and examine the interplay between this subject and health equity concerns.

Improving Internet Skills Among Middle-Aged and Older Residents
This study delves into the usage of the internet among middle-aged and older adults in China, analyzing it through the digital divide framework. As of 2018, around 86% of respondents either demonstrated limited internet skills or lacked access altogether. This figure represents an improvement from the 98% and 97% non–internet use rate observed between 2011 and 2015 [38], indicating a gradual increase in both internet usage and skill acquisition among this demographic. Such a trend suggests an ongoing closure of the first-level digital divide, marked by a shift from nonuse to initial internet engagement. Despite this progress, a second-level digital divide is apparent among those who have embraced the internet and acquired specific skills, with a substantial segment showing advanced internet skills, underscoring an uneven skill distribution. The variability in internet skills may be attributed to individual IT preferences and the consistent support from family and peers in IT usage and learning [20,31,41]. Considering the beneficial impact of internet skills on the mental health of middle-aged and older individuals, and in light of increasing internet penetration rates, it becomes crucial for society and policy makers alike to focus on bridging this second-level digital divide.

IT training programs can reduce anxiety in older adults, increase their interest and efficacy in IT, and improve their IT capability [42]. A community or senior university that provides courses on internet use can follow three guidelines: (1) provide regular long-term training; older adults prefer long-term guidance on IT use and a stable context for experience exchange [42]; (2) leverage the influence of those with good internet skills [43]: interacting with more qualified peers and obtaining guidance from them will reduce the attrition rate of courses; more skilled older adults are important role models and mentors for older adults to improve their skills [43]; and (3) develop web-based participation projects to encourage older adults to establish, maintain, and participate in their own web-based communities [44].

Technical support from the younger generation is vital: young people are “impatient” with older IT learners and lack an understanding of older adults’ special needs [41]. We need to educate younger generations about the benefits of internet use among older adults to inspire families and communities to bridge the digital divide for older adults. Improving internet skills also requires gerontological software design improvements. The Chinese government is actively promoting the gerontological design of internet-based applications [45]. The initial stages of
the project mainly focus on interface design and function simplification. Follow-up studies should develop tools and applications that specifically support older adults’ current activities and goals. Finally, older adults may have negative attitudes toward the internet [31]. It is necessary to reduce the negative connotations surrounding internet use and emphasize that web-based activity can be meaningful and manageable rather than just a form of entertainment.

Limitations
This study has several limitations. First, our findings are based on self-reported data, which implies potential self-reported bias. Second, we assume that the data are missing at random and used multiple imputations to resolve this issue. However, the excluded individuals are more likely to be older, less educated, have a severe disability, and more likely to have higher levels of depressive symptoms. Thus, we cannot exclude this bias. Third, while CHARLS provides a substantial sample size for investigating the relationship between internet skills and depressive symptoms, it is imperative to acknowledge certain shortcomings inherent in the measurement items within the secondary data. This study delineates and quantifies internet skills, differentiated by the level of difficulty associated with various functionalities. Future research should aim to explore more robust measurement methodologies to enhance the reliability of these findings. Lastly, in 2018, a significant majority (98.6%) of Chinese internet users accessed the internet via their mobile phones [46]. Consequently, this study used the rate of mobile phone penetration at the city level as an instrumental variable. This approach was further nuanced by incorporating the percentage of residents aged 60 years and older in each city. Future research should aim to enhance the precision of this instrumental variable by integrating data on city-level internet usage. Additionally, examining the segment of the population aged 45 years and older may offer more relevance, closely mirroring the age range of our study’s sample. Future research can address these issues and explore the formative mechanisms of the second-level digital divide among older adults.

Conclusion
Understanding the role of IT in fulfilling the well-being of older adults has been limited. This study evaluates the effect of internet skills on depressive symptoms through the IV approach. The results reveal the emergence of a second-level digital divide—the disparity in internet skills among middle-aged and older Chinese adults. A 1% improvement in internet skills reduces depressive symptoms by 1.1%-1.7%. This study contributes to the literature on the societal impacts of the internet.

Acknowledgments
This work was supported by the National Natural Science Foundation of China (71874059). The authors thank the participants for their time and contributions to this research.

Authors’ Contributions
AM drafted the main manuscript and ZL reviewed the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Key findings from the literature review, results from the analysis of alternative approaches for assessing internet skills, detailed results of the controls, and detailed results of the ordinary least squares and instrumental variables regressions.

References


Abbreviations

CES-D-10: 10-item Center for Epidemiologic Studies Depression Scale
CFPS: China Family Panel Studies
CHARLS: China Health and Retirement Longitudinal Study
IV: instrumental variables
OLS: ordinary least squares
SES: socioeconomic status

Edited by R Yang; submitted 15.07.23; peer-reviewed by B Wang, W He; revised version received 29.12.23; accepted 20.01.24; published 21.03.24.

Please cite as:
Mu A, Liu Z. Assessing the Impact of Internet Skills on Depressive Symptoms Among Chinese Middle-Aged and Older Adults: Cross-Sectional Instrumental Variables Analysis
JMIR Aging 2024;7:e50880
URL: https://aging.jmir.org/2024/1/e50880
doi: 10.2196/50880

© Aruhan Mu, Zhiyong Liu. Originally published in JMIR Aging (https://aging.jmir.org), 21.3.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Adoption and Use of Telemedicine and Digital Health Services Among Older Adults in Light of the COVID-19 Pandemic: Repeated Cross-Sectional Analysis

Motti Haimi1,2,3, MD, PhD, MHA; Ruslan Sergienko4, MSc

1 Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa, Israel
2 Health Administration Department, The Max Stern Yezreel Valley College, Emek Yezreel, Israel
3 Clalit Research Institute, Tel Aviv, Israel
4 Department of Health Policy and Management, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer Sheva, Israel

Corresponding Author:
Motti Haimi, MD, PhD, MHA
Health Administration Department
The Max Stern Yezreel Valley College
D.N Emek Yezreel
Emek Yezreel, 1930600
Israel
Phone: 972 0504557767
Email: morx@netvision.net.il

Abstract

Background: As the population ages and the prevalence of long-term diseases rises, the use of telecare is becoming increasingly frequent to aid older people.

Objective: This study aims to explore the use and adoption of 3 types of telehealth services among the older population in Israel before, during, and after the COVID-19 pandemic.

Methods: We explored the use characteristics of older adults (aged ≥65 years) belonging to Clalit Health Services in several aspects in the use of 3 types of telehealth services: the use of digital services for administrative tasks; the use of synchronous working-hours telehealth visits with the patient’s personal physician during clinic business hours; and the use of after-hours consultations during evenings, nights, and weekends when the clinics are closed. The data were collected and analyzed throughout 3 distinct periods in Israel: before the COVID-19 pandemic, during the onset of the COVID-19 pandemic, and following the COVID-19 peak.

Results: Data of 618,850 patients who met the inclusion criteria were extracted. Telehealth services used for administrative purposes were the most popular. The most intriguing finding was that the older population significantly increased their use of all types of telehealth services during the COVID-19 pandemic, and in most types, this use decreased after the COVID-19 peak, but to a level that was higher than the baseline level before the COVID-19 pandemic. Before the COVID-19 pandemic, 23.1% (142,936/618,850) of the study population used working-hours telehealth visits, and 2.2% (13,837/618,850) used after-hours consultations at least once. The percentage of use for these services increased during the COVID-19 pandemic to 59.2% (366,566/618,850) and 5% (30,777/618,850) and then decreased during the third period to 39.5% (244,572/618,850) and 2.4% (14,584/618,850), respectively (P<.001). Multiple patient variables have been found to be associated with the use of the different telehealth services in each period.

Conclusions: Despite the limitations and obstacles, the older population uses telehealth services and can increase their use when they are needed. These people can learn how to use digital health services effectively, and they should be given the opportunity to do so by creating suitable and straightforward telehealth solutions tailored to this population and enhancing their usability.

(JMIR Aging 2024;7:e52317) doi:10.2196/52317

KEYWORDS

telemedicine; digital health; older adults; COVID-19; use; digital divide; usability; pandemic; telehealth; Israel; working-hours telehealth visits; after-hours consultation; teleconsultation; eHealth; mobile health; mHealth; wearables; mobile phone
**Introduction**

**Background**

Telehealth is the practice of providing patients with long-distance clinical health care through various communication technologies (television, email, telephone, videoconferencing, internet, and radio) when a patient and physician cannot be present simultaneously [1,2]. Telehealth, a more general term, encompasses health-related education programs such as diabetes management and nutrition seminars. It is distinct from telemedicine, which is more particularly concerned with the delivery of clinical care via the internet [3]. Telemedicine, which uses current information and communication technology, blends patient requirements with technological progress, going beyond the boundaries of traditional health care systems [4].

Telecare combines professional remote health care services with technological tools and assistive technologies. It offers a range of services, including training, monitoring, consultation, communication, and consultation to preserve users' autonomy and improve their quality of life. It is particularly valuable for those who live in remote areas, groups considered vulnerable, and aging populations [1,5].

The world’s population is aging quickly, especially in Europe, according to demographic statistics. Aged populations are now more prevalent than ever in many countries, particularly in the more high-income areas. The percentage of people aged >60 years in the world will double between 2000 and 2050, from approximately 11% to 22%. It is predicted that between 2000 and 2050, the number of persons aged >80 years will double [6]. The aging of the population has resulted in many older people living alone in our communities. Because of instances such as the death of a spouse, older people are increasingly being compelled to singularize or live in a single home [4].

Telecare is used more frequently to assist older people in maintaining their independence and carrying on with their current way of life as the population ages and the prevalence of long-term conditions rises. It appears to be one of the most effective strategies for promoting independent living in a community-dwelling setting because it gives an older person a sense of security and comfort [1,7].

Older people have emerged as one of the primary target demographics for telecare technology in recent years, with a variety of gadgets available for those with long-term medical illnesses as well as for those who have limited mobility or memory issues associated with aging [4,8]. Living at home is associated with a superior quality of life, dignity, and independence, and there is a growing trend among older people to do so rather than age in a health care facility [9,10].

A sizeable portion of the population of older adults have at least 1 chronic illness that necessitates routine monitoring and some level of self-management [11]. However, older patients are less likely to notice signs of an exacerbation before being admitted to the hospital, have less awareness of their disease, and engage in fewer self-management activities [12]. The issue may be resolved by evolving technologies that can notify patients to monitor health status information that can help with at-home self-management [7,13,14]. Although there has been general success for many of the telehealth systems already in use, these technologies are sometimes created without considering how easy they will be for patients and caregivers to use. Although telemedicine offers a way to deliver equitable health care, many people with disabilities find it difficult and challenging to access and use telehealth services [15].

Patients participating in video visits must have the technical knowledge and aptitude to connect to the internet, use and troubleshoot audio-visual equipment, and converse without in-person cues. Due to their limitations or lack of technological skills, many older people might be unable to perform this. In addition, older individuals frequently resist using new technology, especially when it comes to gaining knowledge and learning the skills required to operate computers and other electronic devices [4,16]. Older people may also have changes in their eyesight, hearing, and dexterity in addition to the symptoms of chronic illnesses, which could make it difficult for them to use different telehealth devices [17,18].

Although phone consultations are not ideal for care that necessitates visual assessment, they may increase access for the estimated 6.3 million older people who are unfamiliar with technology or have vision impairment [19]. To safeguard both patients and medical staff during the COVID-19 pandemic, there has been a substantial shift to telemedicine, with video visits being encouraged to see patients at home [3,11].

Telehealth, which allows patients to consult with medical professionals in real time and receive advice on their health issues, has become a basic requirement for the public, especially for those who are in quarantine. Telehealth was the most often used method of service delivery during the pandemic, according to a recent report from the World Health Organization [20]. The study also revealed a trend of rise in telehealth uses as income levels rise; even low-income nations, where 42% of people experienced service interruptions during the COVID-19 pandemic, reported using this technology.

The prevalence of telemedicine unreadiness among Medicare beneficiaries aged ≥65 years in the United States during the COVID-19 pandemic was studied in cross-sectional research in community-dwelling individuals and reported by Lam et al [16]. Patients who met the criteria for unreadiness included those who were older, male, single, Black, or Hispanic; lived in a nonmetropolitan area; and had less education, less income, and worse self-reported health. In total, 72% of adults aged >85 years met those criteria.

Despite the difficulties with using technology mentioned earlier, there is a misperception about older adults that they either lack interest in using technology or are unable to use technological platforms. Contrary to that belief, most older persons (70%) own and regularly use a computer, smartphone, or tablet with an internet connection at home [21]. However, just a small percentage of older people are comfortable using telehealth (11%) [21].
Objectives
Considering the growing phenomena of our aging society and the need to implement telecare for this age group, specifically, this study aimed to explore the use and adoption of 3 types of telehealth services among the older population in Israel before, during, and after the COVID-19 pandemic. Using a quantitative approach, the data have been extracted before the COVID-19 pandemic, throughout the pandemic, and during the months after the peak of the epidemic in Israel.

In addition, we wanted to determine whether the COVID-19 pandemic had an impact on how older adults used telehealth services and, if so, whether those changes would last once the pandemic concluded. The results of this study will enable us to emphasize to health care decision makers the necessity for tailored telemedicine-based care that considers the needs, abilities, and preferences of the older population and adapts over time as those needs change.

Methods

Study Population
Clalit Health Services, the largest integrated health care service provider and payer system in Israel, has >4 million active members. Clalit Health Services has a comprehensive health care data warehouse, which integrates hospital and community medical records, laboratory and imaging information, pharmaceutical records, health care expenses, and Ministry of the Interior vital statistics of all the members. Clalit Health Services experiences membership turnover of <1% annually, making it easier to track population trends over time. The inclusion and exclusion criteria are presented inTextbox 1.

Textbox 1. Inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Membership in Clalit Health Services for at least a year</td>
</tr>
<tr>
<td>• Aged ≥65 years at each period</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>• No Clalit Health Services membership</td>
</tr>
<tr>
<td>• Aged &lt;65 years at each period</td>
</tr>
</tbody>
</table>

Ethics Approval
The study was ethically approved by the boards of the Clalit Health Services on January 18, 2021 (reference numbers 826 and COM-0113-21).

Study Design
This study is a repeated cross-sectional analysis. This type of study looks at data collected at a single point in time, rather than over a period, which is useful for comparing and analyzing the effect of different factors on one another or describing a sample.

Data Extraction
We explored the telehealth use characteristics of older adults (aged ≥65 years) belonging to Clalit Health Services in several aspects:

1. The use of digital services for administrative tasks such as web-based medical appointment scheduling and physician request submission. This category of eHealth services was named administrative.
2. The use of synchronous web-based, telephone, or digital visits with the patient’s personal physician during clinic business hours (through video or telephone), initiated by the patients themselves. This type of consultation was named working hours telehealth consultations.
3. The use of web-based or telephone consultations (not with the personal physician) during the evenings, nights, and weekends when the clinics are closed, including the use of a phone, video camera, or “tytocare” equipment [22]. The term after-hours consultations was used to describe this form of consultation.

Data Extraction Periods
Three separate periods were used to collect and analyze the data:

1. Baseline—from February 1, 2019, to the end of February 2020, the period before the COVID-19 pandemic in Israel, which was labeled period 1 (before).
2. Initiation—from March 1, 2020, to March 31, 2021, when lockdowns were implemented during the COVID-19 pandemic in Israel. Period 2 (during) was used to designate this time frame.
3. Follow-up—from April 1, 2021, to the end of October 2021, following the COVID-19 pandemic peak in Israel. Period 3 (after) was used to designate this time frame.

Study Timeline
Figure 1 describes the timing scheme. The “index date period” refers to the time from March 1, 2020, to March 31, 2021, which is considered the peak of COVID-19 pandemic in Israel. The study was designed and planned during this period.
We planned to gather certain data retrospectively from February 1, 2019, to the index date period and the rest of the data prospectively from the index date period to the end of the follow-up period date, which was October 31, 2021. As intended, and previously indicated, the participant files were recruited from February 1, 2019, to October 31, 2021.

Finally, after gaining approval from all committees of Clalit Health Services, the raw data were accessed for research purposes between February 20, 2022, and May 16, 2022.

Study Outcomes
As previously indicated, the main outcomes (dependent variables) measured were as follows: (1) The use of digital services for administrative tasks, such as web-based medical appointment scheduling and physician request submission, named administrative services (2) The use of video or telephone calls with a personal physician conducted during clinic business hours at the patients’ request, named working-hours telehealth consultations (3) The use of web-based consultations (not with the personal physician) during the evenings, nights, and weekends when the clinics are closed, including the use of the telephone, video camera, or Tytocare device, named after-hours consultations.

Covariables (Covariates)
The inpatient and outpatient data warehouses of Clalit Health Services have provided the data. The information covered the use of telemedicine and digital health services during the aforementioned 3 periods. The variables can be described as dichotomous (use of telemedicine and digital health; yes or no) and as categorical in the aspect of which digital services were used: (1) administrative only use; (2) consultations with the personal physician—working-hours telehealth consultations; (3) consultations during the after hours of the clinic—after-hours consultations.

Multiple patient variables, such as sociodemographic information (age, sex, place of birth, place of residence, socioeconomic status [SES], ethnicity, country of birth, etc), clinical information (chronic diseases, habits, etc), and use of chronic medications have been examined for each period.

The different features were extracted and categorized as follows: (1) sociodemographic parameters, including sex, age, SES, country of birth (coalesced into regions when necessary), ethnicity by country of individual’s or parents’ birth, sector (clinic level data—predominantly Arab or Jewish), marital status and number of children, Clalit Health Services affiliation by district, subdistrict, and clinic; (2) clinical markers or comorbidities, such as smoking status, alcohol use (and related diagnoses), BMI, height, weight, Charlson comorbidity index (the most widely used comorbidity index used to determine survival rate [1 year and 10 year] in patients with multiple comorbidities), presence of chronic diseases such as active malignancy, cardiovascular diseases (ischemic heart disease, cerebrovascular disease, hypertension), asthma, diabetes, neurological diseases (Alzheimer disease and Parkinson disease), psychiatric disease, diagnosis of COVID-19 since the index period; (3) use of chronic medications, especially antianxiety, antidepressant, and sleep-aid medications.

Statistical Analysis
We analyzed the data according to the type of telehealth use (working hours telehealth consultations, after-hours consultations, and administrative), the time frame in which it was used (before, during, or after the COVID-19 pandemic in Israel), and the amount of use (how often it was used) according to 2 categories (0 or ≥1 time).

We used appropriate descriptive statistics to characterize the study population. The association between telemedicine use and each available socioeconomic factor was studied using

---

**Figure 1.** Study design timeline.

<table>
<thead>
<tr>
<th>Retroactive</th>
<th>Follow-up (prospective)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrospective</td>
<td>Follow-up (prospective)</td>
</tr>
<tr>
<td>After the COVID-19 “initiation”: April 2021 to October 2021</td>
<td></td>
</tr>
</tbody>
</table>
univariable analyses while comparing participants who used telemedicine at least once during the specific period with patients with zero use (Pearson $\chi^2$ test). We have used nonparametric related samples Cochran $Q$ test to compare telemedicine use during the 3 periods.

Post hoc analysis with the Dunn and McNemar post hoc tests was conducted (with a Bonferroni correction applied) to access both between-subjects and within-subjects effects, analyzing each pair of periods. In addition, we performed a set of multivariable binary logistic regression models to estimate the association between telemedicine use and socioeconomic factors during each period. These models were used to calculate odd ratios (ORs) and 95% CIs. A $P$ value <.05 was considered statistically significant. SPSS IBM Statistics for Windows, predictive analytics software (version 28.0 [28.0.1]), was used for data analysis.

Privacy

Data extraction was conducted by the research room team at Clalit Health Services. The deidentified, raw extracted data were stored on the virtualization desktop infrastructure (VDI), a secure setting. Data were analyzed on the VDI, and only aggregated nonidentifiable results were moved out of the VDI for publication.

As discussed earlier, we did not have any access to identified information. We had limited and restricted access only to unidentified data. A confirmation from the special committee for data mining from Clalit Health Services authorities was received on August 25, 2021.

Results

General Characteristics

There were 669,349 patients in total who met the inclusion criteria at the start of period 1 (2019). In 2020, a total of 642,223 patients met the criterion for inclusion, whereas in 2021, only 618,850 patients met the requirements. Various analyses were carried out on this population. The general characteristics are presented in Table 1.

The mean Charlson score was 5.59 (SD 2.57), with a median of 5.00 (range 2-22). The mean age in 2019 was 75.16 (SD 7.64) years, with a median of 73.00 (range 65-110) years.
<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Patients, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>349,069 (56.4)</td>
</tr>
<tr>
<td>Male</td>
<td>269,781 (46.6)</td>
</tr>
<tr>
<td><strong>Age group in 2019 (years)</strong></td>
<td></td>
</tr>
<tr>
<td>65-74</td>
<td>349,477 (56.5)</td>
</tr>
<tr>
<td>75-84</td>
<td>188,110 (30.4)</td>
</tr>
<tr>
<td>≥85</td>
<td>81,263 (13.1)</td>
</tr>
<tr>
<td><strong>Country of birth</strong></td>
<td></td>
</tr>
<tr>
<td>Israel</td>
<td>231,207 (37.4)</td>
</tr>
<tr>
<td>Other</td>
<td>387,643 (62.6)</td>
</tr>
<tr>
<td><strong>Socioeconomic status</strong></td>
<td></td>
</tr>
<tr>
<td>Low</td>
<td>146,469 (23.7)</td>
</tr>
<tr>
<td>Medium</td>
<td>235,231 (38)</td>
</tr>
<tr>
<td>High</td>
<td>237,150 (38.3)</td>
</tr>
<tr>
<td><strong>Demographic sector</strong></td>
<td></td>
</tr>
<tr>
<td>General Jewish</td>
<td>527,349 (85.2)</td>
</tr>
<tr>
<td>Cherkess</td>
<td>355 (0.1)</td>
</tr>
<tr>
<td>Religious Jewish</td>
<td>12,134 (2)</td>
</tr>
<tr>
<td>Arab</td>
<td>56,692 (9.2)</td>
</tr>
<tr>
<td>Unknown</td>
<td>22,320 (3.6)</td>
</tr>
<tr>
<td><strong>District</strong></td>
<td></td>
</tr>
<tr>
<td>South</td>
<td>64,985 (10.5)</td>
</tr>
<tr>
<td>Center</td>
<td>291,145 (47)</td>
</tr>
<tr>
<td>North</td>
<td>168,783 (27.3)</td>
</tr>
<tr>
<td>Center east</td>
<td>93,937 (15.2)</td>
</tr>
<tr>
<td><strong>Smoking status</strong></td>
<td></td>
</tr>
<tr>
<td>Never smoked</td>
<td>391,918 (63.3)</td>
</tr>
<tr>
<td>Past smoker</td>
<td>126,853 (20.5)</td>
</tr>
<tr>
<td>Current smoker</td>
<td>61,275 (9.9)</td>
</tr>
<tr>
<td>Status unknown</td>
<td>38,804 (6.3)</td>
</tr>
<tr>
<td><strong>Any chronic disease</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>520,771 (84.2)</td>
</tr>
<tr>
<td>No</td>
<td>98,079 (15.8)</td>
</tr>
<tr>
<td><strong>Active malignancy</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>57,452 (9.3)</td>
</tr>
<tr>
<td>No</td>
<td>561,398 (90.7)</td>
</tr>
<tr>
<td><strong>Cardiovascular disease</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>479,758 (77.5)</td>
</tr>
<tr>
<td>No</td>
<td>139,092 (22.5)</td>
</tr>
<tr>
<td><strong>Asthma</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>109,553 (17.7)</td>
</tr>
<tr>
<td>No</td>
<td>509,297 (82.3)</td>
</tr>
</tbody>
</table>
Patients, n (%)  
Characteristics | Diabetes | 237,968 (38.5) | 380,882 (61.5) |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>237,968 (38.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>380,882 (61.5)</td>
<td></td>
</tr>
</tbody>
</table>
Neurological disease  
Yes | 37,814 (6.1) |
No | 581,036 (93.9) |

**Telehealth Use Characteristics**

**Overview**

Table 2 lists the visit counts (per person) according to the telehealth use type (administrative, working hours, or after hours) and time frame. For telehealth administrative purposes, we can observe that the mean count (per person) was 4.39 (SD 5.56) before the COVID-19 pandemic (period 1), increased to 5.55 (SD 6.71) during the COVID-19 pandemic (period 2), and decreased to 3.07 (SD 4.11) after the peak of the COVID-19 pandemic in Israel (period 3).

Regarding working-hours telehealth consultations with the personal physician, the mean number (per person) was 0.49 (SD 1.43) before the COVID-19 pandemic (period 1), it increased dramatically to a mean of 2.23 (3.54) during the COVID-19 pandemic (period 2), and then it decreased to a mean of 1.00 (1.95) after the peak of the COVID-19 pandemic in Israel (period 3), but still higher than that in period 1.

The mean after-hours telehealth use (per person) was 0.03 (SD 0.02) before the COVID-19 pandemic, increased to 0.07 (0.46) during the COVID-19 pandemic, and decreased again after the peak of the COVID-19 pandemic to 0.03 (SD 0.33).

The visit counts were also analyzed in a dichotomic manner into 2 categories: “no” or “yes,” that is, 0 visits versus ≥1 visit, respectively.

The comparison of telemedicine use (at least once) during 3 periods and the results of Cochran Q test are presented in Figure 2 and Table 3, respectively. The pairwise comparisons of telemedicine use (at least once) during the 3 periods, results of the Dunn and McNemar post hoc tests (with Bonferroni correction), between and within subjects, respectively, are presented in Tables 4 and 5.

**Table 2.** Telehealth use (visit counts) according to types and periods.

<table>
<thead>
<tr>
<th>Telehealth use</th>
<th>Visits, mean (SD)</th>
<th>Visits, median (IQR)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Period 1 (before)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrative</td>
<td>4.39 (5.56)</td>
<td>3.00 (0-191)</td>
</tr>
<tr>
<td>Working hours</td>
<td>0.49 (1.43)</td>
<td>0.00 (0-66)</td>
</tr>
<tr>
<td>After hours</td>
<td>0.03 (0.02)</td>
<td>0.00 (0-83)</td>
</tr>
<tr>
<td><strong>Period 2 (during)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrative</td>
<td>5.55 (6.71)</td>
<td>4.00 (0-208)</td>
</tr>
<tr>
<td>Working hours</td>
<td>2.23 (3.54)</td>
<td>1.00 (0-104)</td>
</tr>
<tr>
<td>After hours</td>
<td>0.07 (0.46)</td>
<td>0.00 (0-96)</td>
</tr>
<tr>
<td><strong>Period 3 (after)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrative</td>
<td>3.07 (4.11)</td>
<td>2.00 (0-107)</td>
</tr>
<tr>
<td>Working hours</td>
<td>1.00 (1.95)</td>
<td>0.00 (0-50)</td>
</tr>
<tr>
<td>After hours</td>
<td>0.03 (0.33)</td>
<td>0.00 (0-96)</td>
</tr>
</tbody>
</table>
**Working-Hours Telehealth Visits (During the Regular Work Hours, With the Personal Physician)**

During the first period, 23.1% (142,936/618,850) of the study population used working-hours telehealth services at least once. The percentage of use increased to 59.2% (366,566/618,850) during the second period and then decreased to 39.5% (244,572/618,850) during the third period (still higher than that in period 1).

Pairwise comparisons were performed using the Dunn [23] procedure (between-subjects analysis) with a Bonferroni correction for multiple comparisons. Adjusted P values are presented in Table 4. Compared to the percentage of working-hours telehealth services use during the first period, there was a statistically significant increase in the percentage of use during the second period (P<.001) and during the third period (P<.001). There was also a statistically significant decrease in the percentage of the study population that used working-hours telehealth services at least once during the third period compared to the second period (P<.001).
Similar results were demonstrated using the McNemar post hoc test (within-subjects analysis with Bonferroni correction), as demonstrated in Table 5.

**After-Hours Telemedicine**

During the first period, 2.2% (13,837/618,850) of the study population used after-hours telemedicine services at least once. The percentage of use increased to 5.0% (30,777/618,850) during the second period and then decreased to 2.4% (14,584/618,850) during the third period (still higher than that in period 1).

Pairwise comparisons were performed using the Dunn [23] procedure (between-subjects analysis) with a Bonferroni correction for multiple comparisons. Adjusted \( P \) values are presented in Table 4. Compared to the percentage of after-hours services use during the first period, there was a statistically significant increase in the percentage of use during the second period \( (P<.001) \), while there was no statistically significant difference compared to the third period. There was also a statistically significant decrease in the percentage of the study population that used after-hours telemedicine services at least once during the third period compared to the second period \( (P<.001) \).

When using the McNemar post hoc test (within-subjects analysis with Bonferroni correction), significant differences were demonstrated between all 3 pairs of periods, as demonstrated in Table 5.

**Administrative Telemedicine**

During the first period, 69% (427,295/618,850) of the study population used administrative telemedicine services at least once. The percentage of use increased to 74.3% (459,622/618,850) during the second period and then decreased to 67.9% (420,209/618,850) during the third period. Pairwise comparisons were performed using the Dunn [23] procedure (between-subjects analysis) with a Bonferroni correction for multiple comparisons. Adjusted \( P \) values are presented in Table 4. Compared to the percentage of administrative services use during the first period, there was a statistically significant increase in the percentage of use during the second period \( (P<.001) \), while there was no statistically significant difference compared to the third period. There was also a statistically significant decrease in the percentage of the study population that used administrative telemedicine services at least once during the third period compared to the second period \( (P<.001) \).

When using the McNemar post hoc test (within-subjects analysis with Bonferroni correction), significant differences were demonstrated between all 3 pairs of periods, as demonstrated in Table 5.

**Univariate Analysis of Working-Hours Telehealth Visits**

According to the dichotomous classification into 2 categories (0 uses versus \( \geq 1 \) uses), a univariate analysis was conducted using Pearson \( \chi^2 \) tests on the various types of telehealth use, and the results are shown in Multimedia Appendices 1-3. Multimedia Appendix 1 demonstrates the significant factors associated with the working-hours telehealth visits (telehealth services with the personal physician), with the 2 categories classification (0 visits vs \( \geq 1 \) visits).

**Gender**

A higher percentage of female individuals than male individuals used the web-based visits at all periods. In addition, both male individuals and female individuals increased their working-hours telecare use during the COVID-19 pandemic period, which was followed by a decline, though at a higher level than before (period 1).

**Age Group**

The age group of 75 to 84 years had used the web-based services to a greater extent, compared to the other age groups (65-74 years and \( >85 \) years), at all the 3 periods. In addition, during the COVID-19 pandemic, all age groups significantly boosted their use of working-hours telehealth services; following the COVID-19 pandemic, the use reduced but remained significantly greater than it was before the COVID-19 pandemic (period 1).

**Country of Birth**

Before the COVID-19 pandemic, persons who were born in Israel used working-hours telehealth services more frequently than those who were born elsewhere. However, this tendency shifted during and after the COVID-19 pandemic, when those born outside of Israel had a larger use. Both groups had increased their use of working-hours telehealth services during the COVID-19 pandemic, and it had declined after that time to a greater level than it had been before the COVID-19 pandemic.

**SES Level**

Persons belonging to a higher SES level had a higher level of working-hours telehealth services use, compared to persons with lower SES, at all periods. Here also, we see that during the COVID-19 pandemic, persons in all SES levels greatly expanded their use of working-hours telehealth services; this use eventually declined, but it remained higher than it had been before the COVID-19 pandemic.

**Demographic Sector**

If we ignore the “unknown” portion (which accounts for only 3.6% of the populations included), we see that the religious Jewish population (followed by the general Jewish population) had higher use of working hours telehealth services, as compared to other sectors, at all periods. Here again, persons in all demographic sectors greatly expanded their use of working hours telehealth services during the COVID-19 pandemic period; this use eventually declined, but it remained higher than it had been before the COVID-19 pandemic.

**District**

People from the northern region of Israel had more working hours telehealth visits than people from other districts before the COVID-19 pandemic (period 1). Those from the center districts, however, made many more working-hours telehealth visits to their physician both during and after the COVID-19 pandemic. As stated in the factors described earlier, all subgroups had increased their use of working-hours telehealth services during the COVID-19 pandemic period, and it had
declined after that time to a greater level than it had been before the COVID-19 pandemic.

**Smoking Status**

Those who were past smokers used the working hours telehealth visits more frequently than people who are currently smokers or even people who had never smoked. This held true throughout all periods. All subgroups had increased their use of working-hours telehealth services during the COVID-19 pandemic, and it had decreased after that time to a greater degree than it had been before the COVID-19 pandemic, as reflected in the factors mentioned earlier.

**The Presence of a Chronic Disease**

Across all periods, those with chronic diseases significantly used working-hours telehealth visits to a greater extent (more than twice as often as people without chronic diseases). Both groups had expanded their use of working hours telehealth services during the COVID-19 pandemic, and use declined after that period to a larger extent than it had before (period 1).

**Univariate Analysis of Administrative Telehealth Use**

Multimedia Appendix 2 shows a univariate analysis, outlining the important variables that were significantly linked to administrative telehealth use in accordance with the 2 categories used (0 visits vs ≥1 visits).

**Sex**

At all 3 periods, a greater proportion of female individuals than male individuals used the administrative web-based services. In contrast to the situation outlined for web-based visits, the administrative telehealth use (for this parameter and all other parameters described here) did increase throughout the COVID-19 pandemic before declining to a level below that of the before period (period 1).

**Age Group**

The age group of 75 to 84 years had used the administrative web-based services in a greater extent, compared to the other age groups (65-74 years and >85 years), at all the 3 periods. In addition, as reported earlier, during the COVID-19 pandemic, all age groups significantly boosted their use of web-based services, but the use reduced following the COVID-19 pandemic, to a lower level than it was before the COVID-19 pandemic.

**Country of Birth**

Throughout all periods, those born outside of Israel used administrative web-based services more than those who were born there. Both groups had increased their use of web-based services during the COVID-19 pandemic, and it had declined after that time to a lower level than it had been before the COVID-19 pandemic.

**SES Level**

Persons belonging to a higher SES level had a higher level of administrative web-based services use, compared to persons with lower SES, at all periods. Here also, we see that during the COVID-19 pandemic, persons in all SES levels greatly expanded their use of web-based services; this use eventually declined to a lower level than it had been before the COVID-19 pandemic.

**Demographic Sector**

Throughout all periods, the Cherkess (Circassians) had the higher use of administrative web-based services, as compared to other sectors. In this instance, as well, people in all demographic groups significantly increased their use of administrative web-based services during the COVID-19 pandemic; nevertheless, this use gradually decreased to a level that was lower than it had been before the COVID-19 pandemic.

**District**

People from Israel’s southern region used administrative internet services to a greater extent than residents from other districts at all times. All subgroups had increased their use of administrative web-based services during the COVID-19 pandemic, and it had decreased after that time to a lower level than it had been before the COVID-19 pandemic, as was shown in other factors mentioned earlier.

**Smoking Status**

People who had previously smoked used the administrative web-based services more than those who smoke now or even those who have never smoked. This was accurate at all times. All subgroups had grown their use of administrative web-based services during the COVID-19 pandemic and had declined after that period to a lesser extent than it had been before the COVID-19 pandemic.

**The Presence of a Chronic Disease**

Throughout all periods, persons with chronic conditions significantly used the administrative web-based services more frequently (more than twice as often as people without chronic diseases). Both groups had expanded their use of web-based services during the COVID-19 pandemic, and use decreased after that period to a lower extent than it had before (period 1).

As demonstrated in Table 2, the overall use of after-hours telemedicine services (during the evenings, nights, and weekend days) was considerably lower than that of administrative telehealth and working-hours telemedicine services (with the personal physician during opening hours).

**Univariate Analysis of After-Hours Telemedicine Use**

In accordance with the 2 categories used (0 visits vs ≥1 visits), Multimedia Appendix 3 presents a univariate analysis summarizing the relevant variables that were significantly associated to after-hours telemedicine use:

**Sex**

At all 3 periods, a greater proportion of females than males used the after-hours telemedicine services. During the COVID-19 pandemic, both men and women increased their use of after-hours telemedicine services, which was followed by a fall, but at a greater level than previously (period 1), in a manner similar to the trend outlined for the working-hours telehealth services.
Age Group

Only on the COVID-19 pandemic, the age groups of 65 to 74 years and 75 to 84 years had significantly greater use of after-hours telemedicine services, compared to the third age group (>85 years). In addition, as previously mentioned, during the COVID-19 pandemic, all age groups increased their use of after-hours services; however, after the COVID-19 pandemic, use decreased, though it remained at a higher level than it had been before the COVID-19 pandemic.

Country of Birth

During the COVID-19 pandemic and the period after the COVID-19 pandemic, those born outside of Israel significantly used after-hours telemedicine services more than those who were born there. Both groups had increased their use of after-hours telemedicine services during the COVID-19 pandemic, and it had declined after that time (for people born outside Israel, it declined to a higher level than it was before the COVID-19 pandemic).

SES Level

Similar to how it was with the other telehealth services, those with higher SES levels used after-hours telemedicine services more frequently than people with lower SES levels did always. In this area as well, the use of after-hours services by people of all SES levels significantly increased during the COVID-19 pandemic. Eventually, this use decreased but it did so at a greater level than it had been before the COVID-19 pandemic.

Demographic Sector

As described for the use of working-hours telehealth visits, here also, the religious Jewish population and the general Jewish population had the higher use of after-hours services, as compared to other sectors, at all periods (the religious Jews had a higher use during the COVID-19 pandemic). Here again, people in all demographic sectors greatly expanded their use of after-hours services during the COVID-19 pandemic, which declined after that period.

District

Contrary to the previously mentioned telehealth services, persons from Israel’s central area were much more likely to use after-hours telemedicine services than those from other regions during all periods. The use of after-hours internet services had surged across all subgroups during the COVID-19 pandemic and had fallen thereafter to a greater level than it had been before the COVID-19 pandemic (except for the south region in which it declined to the same level as before).

Smoking Status

As was the case with the other telehealth services, people who had previously smoked used more the after-hours services than those who do so now or even those who have never smoked. This was accurate at all periods. All groupings had increased their use of after-hours internet services during the COVID-19 pandemic and had decreased after that typically to a greater extent than it had been before the COVID-19 pandemic.

The Presence of a Chronic Disease

Similar to other telehealth services mentioned earlier, those with chronic conditions significantly used the after-hours web-based services more often across all periods (more than twice as often as people without chronic diseases). During the COVID-19 pandemic, both groups (with and without chronic diseases) increased their use of web-based services, and after that time, use decreased but to a higher level than it had previously (period 1).

Multivariate Analysis

In the multivariate analysis, using a set of multivariable binary logistic regression models, several parameters were found to be significantly associated with the different types of telehealth use during each period: male sex (decreased use of all types of telehealth services, during all periods); country of birth–Israel (increased use of working-hours telehealth visits and after-hours visits during all periods but decreased administrative use at all periods); presence of any chronic disease (highly increased use of all types of telehealth services during all periods); Charlson comorbidity score (increased use of all types of telehealth services during all periods); medium and high SES (increased use of all types of telehealth services during all periods); Jewish religion (highly increased use of all types of telehealth services during all periods); southern district—place of residence (people living in this district used fewer working-hours telehealth services and fewer after-hours services during all periods but used more administrative services before and after the COVID-19 pandemic); northern district (people living here used more working-hours telehealth services before the COVID-19 pandemic but used less working-hours telehealth services during periods 2 and 3 and used fewer administrative services and fewer after-hours services at all periods); central east district (people living here used fewer of all types of telehealth services during all periods); current or past smokers (people used more working-hours telehealth services, more after-hours services, and more administrative services during all periods); older age groups of those aged ≥75 years (used fewer working-hours telehealth services, fewer after-hours services, and fewer administrative services during all periods).

The multivariate analysis is demonstrated in Table 6 (for working-hours telehealth visits), Table 7 (for after-hours visits), and Table 8 (for administrative uses). The Nagelkerke $R^2$ values for each type of telehealth service and each period are also provided in the tables.
Table 6. Multivariate analysis—working-hours telehealth visits. Variables entered on step 1: sex (male), country of birth, any chronic disease, Charlson score, SES (medium), SES (high), Jewish nationality, district (south), district (north), district (central east), current or past smoker, age group (75-84 years), age group (≥85 years).

<table>
<thead>
<tr>
<th>Variables in the equation</th>
<th>Working-hours visits</th>
<th>Before (period 1)</th>
<th>During (period 2)</th>
<th>After (period 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Exp (B)</td>
<td>95% CI</td>
<td>Exp (B)</td>
</tr>
<tr>
<td>95% CI</td>
<td>P value</td>
<td></td>
<td></td>
<td>P value</td>
</tr>
<tr>
<td>Sex (male)</td>
<td>&lt;.001</td>
<td>0.717</td>
<td>0.708-0.726</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Country of birth (Israel)</td>
<td>&lt;.001</td>
<td>1.079</td>
<td>1.064-1.094</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Any chronic disease</td>
<td>&lt;.001</td>
<td>1.978</td>
<td>1.934-2.022</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Charlson score</td>
<td>&lt;.001</td>
<td>1.144</td>
<td>1.141-1.147</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>SES (medium)</td>
<td>&lt;.001</td>
<td>1.332</td>
<td>1.308-1.355</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>SES (high)</td>
<td>&lt;.001</td>
<td>1.374</td>
<td>1.349-1.400</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Jewish religion</td>
<td>&lt;.001</td>
<td>1.065</td>
<td>1.042-1.088</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>District (south)</td>
<td>&lt;.001</td>
<td>0.653</td>
<td>0.638-0.668</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>District (north)</td>
<td>&lt;.001</td>
<td>1.109</td>
<td>1.092-1.125</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>District (central east)</td>
<td>.01</td>
<td>0.977</td>
<td>0.960-0.995</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Current or past smoker</td>
<td>&lt;.001</td>
<td>1.152</td>
<td>1.136-1.168</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age group (75-84 years)</td>
<td>&lt;.001</td>
<td>0.818</td>
<td>0.806-0.830</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age group (≥85 years)</td>
<td>&lt;.001</td>
<td>0.743</td>
<td>0.728-0.758</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Constant</td>
<td>&lt;.001</td>
<td>0.068</td>
<td>N/A</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

a Nagelkerke $R^2=0.065$.
b Nagelkerke $R^2=0.139$.
c Nagelkerke $R^2=0.090$.
d SES: socioeconomic status.
e N/A: not applicable.
Table 7. Multivariate analysis—after-hours telehealth visits. Variables entered on step 1: sex (male), country of birth, any chronic disease, Charlson score, SES (medium), SES (high), Jewish nationality, district (south), district (north), district (central east), current or past smoker, age group (75-84 years), age group (≥85 years).

<table>
<thead>
<tr>
<th>After-hours visits</th>
<th>Variables in the equation</th>
<th>Before (period 1)(^a)</th>
<th>During (period 2)(^b)</th>
<th>After (period 3)(^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>P value</td>
<td>Exp (B)</td>
<td>95% CI</td>
</tr>
<tr>
<td>Sex (male)</td>
<td></td>
<td>&lt;.001</td>
<td>0.748</td>
<td>0.722-0.775</td>
</tr>
<tr>
<td>Country of birth (Israel)</td>
<td></td>
<td>&lt;.001</td>
<td>1.086</td>
<td>1.046-1.128</td>
</tr>
<tr>
<td>Any chronic disease</td>
<td></td>
<td>&lt;.001</td>
<td>2.063</td>
<td>1.925-2.210</td>
</tr>
<tr>
<td>Charlson score</td>
<td></td>
<td>&lt;.001</td>
<td>1.111</td>
<td>1.103-1.118</td>
</tr>
<tr>
<td>SES(^d) (medium)</td>
<td></td>
<td>&lt;.001</td>
<td>1.264</td>
<td>1.199-1.333</td>
</tr>
<tr>
<td>SES (high)</td>
<td></td>
<td>&lt;.001</td>
<td>1.404</td>
<td>1.330-1.482</td>
</tr>
<tr>
<td>Jewish religion</td>
<td></td>
<td>&lt;.001</td>
<td>1.981</td>
<td>1.837-2.136</td>
</tr>
<tr>
<td>District (south)</td>
<td></td>
<td>&lt;.001</td>
<td>0.792</td>
<td>0.745-0.842</td>
</tr>
<tr>
<td>District (north)</td>
<td></td>
<td>&lt;.001</td>
<td>0.767</td>
<td>0.734-0.801</td>
</tr>
<tr>
<td>District (central east)</td>
<td></td>
<td>&lt;.001</td>
<td>0.850</td>
<td>0.809-0.893</td>
</tr>
<tr>
<td>Current or past smoker</td>
<td></td>
<td>.001</td>
<td>1.067</td>
<td>1.028-1.108</td>
</tr>
<tr>
<td>Age group (75-84 years)</td>
<td></td>
<td>&lt;.001</td>
<td>0.780</td>
<td>0.749-0.812</td>
</tr>
<tr>
<td>Age group (≥85 years)</td>
<td></td>
<td>&lt;.001</td>
<td>0.730</td>
<td>0.691-0.772</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>.000</td>
<td>0.004</td>
<td>N/A(^e)</td>
</tr>
</tbody>
</table>

\(^a\)Nagelkerke \(R^2=0.027\).
\(^b\)Nagelkerke \(R^2=0.027\).
\(^c\)Nagelkerke \(R^2=0.022\).
\(^d\)SES: socioeconomic status.
\(^e\)N/A: not applicable.
Table 8. Multivariate analysis—administrative telehealth visits. Variables entered on step 1: sex (male), country of birth, any chronic disease, Charlson score, SES (medium), SES (high), Jewish nationality, district (south), district (north), district (central east), current or past smoker, age group (75-84 years), age group (≥85 years).

<table>
<thead>
<tr>
<th>Administrative telehealth visits</th>
<th>Variables in the equation</th>
<th>Before (period 1)</th>
<th>During (period 2)</th>
<th>After (period 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>( P ) value</td>
<td>( \text{Exp (B)} )</td>
<td>( 95% \text{ CI} )</td>
</tr>
<tr>
<td>Sex (male)</td>
<td></td>
<td>&lt;.001</td>
<td>0.788</td>
<td>0.778-0.798</td>
</tr>
<tr>
<td>Country of birth (Israel)</td>
<td></td>
<td>&lt;.001</td>
<td>0.938</td>
<td>0.926-0.951</td>
</tr>
<tr>
<td>Charlson score</td>
<td></td>
<td>&lt;.001</td>
<td>1.213</td>
<td>1.209-1.216</td>
</tr>
<tr>
<td>SES (medium)</td>
<td></td>
<td>&lt;.001</td>
<td>1.355</td>
<td>1.333-1.377</td>
</tr>
<tr>
<td>SES (high)</td>
<td></td>
<td>&lt;.001</td>
<td>1.386</td>
<td>1.362-1.410</td>
</tr>
<tr>
<td>Jewish religion</td>
<td></td>
<td>&lt;.001</td>
<td>2.291</td>
<td>2.247-2.335</td>
</tr>
<tr>
<td>District (south)</td>
<td></td>
<td>&lt;.001</td>
<td>1.044</td>
<td>1.023-1.067</td>
</tr>
<tr>
<td>District (north)</td>
<td></td>
<td>&lt;.001</td>
<td>0.655</td>
<td>0.646-0.665</td>
</tr>
<tr>
<td>District (central east)</td>
<td></td>
<td>&lt;.001</td>
<td>0.702</td>
<td>0.690-0.714</td>
</tr>
<tr>
<td>Current or past smoker</td>
<td></td>
<td>&lt;.001</td>
<td>1.311</td>
<td>1.293-1.329</td>
</tr>
<tr>
<td>Age group (75-84 years)</td>
<td></td>
<td>&lt;.001</td>
<td>0.753</td>
<td>0.743-0.764</td>
</tr>
<tr>
<td>Age group (≥85 years)</td>
<td></td>
<td>&lt;.001</td>
<td>0.452</td>
<td>0.444-0.461</td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>&lt;.001</td>
<td>0.162</td>
<td>N/A</td>
</tr>
</tbody>
</table>

\[a\text{Nagelkerke } R^2=0.198.\]
\[b\text{Nagelkerke } R^2=0.209.\]
\[c\text{Nagelkerke } R^2=0.193.\]
\[d\text{SES: socioeconomic status.}\]
\[e\text{N/A: not applicable.}\]

**Discussion**

**Principal Findings**

In this study, we examined the use and uptake of 3 distinct telehealth services among the older population who were members of Clalit Health Services before, during, and after the COVID-19 pandemic in Israel. Data of 618,850 patients who met the inclusion criteria were extracted. Telehealth services used for administrative purposes were the most popular. The older population significantly increased their use of all types of telehealth services during the COVID-19 pandemic, and in most types of services, this use decreased after the COVID-19 peak but to a level that was higher than the baseline level before the COVID-19 pandemic. The three telehealth service types that were investigated in the study were as follows: (1) “working-hours telehealth visits” refers to video or telephone consultations with a personal physician during clinic business hours; (2) “administrative” refers to the use of digital services for administrative tasks (such as scheduling appointments or submitting requests to the physicians on the web); and (3) “after-hours visits” refers to the use of web-based consultations in the evenings, weekend days, and other nonbusiness hours (with other physicians). A quantitative method was used to extract the data at three different time points: (1) “before” was before the COVID-19 pandemic; (2) “during” was during the COVID-19 pandemic’s emergence; and (3) “after” was during the months after the peak of the pandemic in Israel.

Our main objectives were to assess how the older adults in Israel used various telehealth services, considering the challenges and difficulties they encountered; to determine what factors were associated with increased or decreased uptake; and to determine whether the COVID-19 period had any impact on use patterns and whether those patterns persisted after the period had ended. Tables 2 and 3 show that, throughout the study periods, the telehealth services used for administrative purposes were the most popular among the older population. This was followed by “working-hours telehealth visits”—telemedicine consultations with the personal physician—during the regular business hours of clinics. Among the other services mentioned earlier, “after-hours” telemedicine visits came in last.
Another intriguing finding was that during the COVID-19 pandemic, the older population significantly increased their use of all telehealth services, and this use decreased after the peak of the COVID-19 pandemic (period 3), in all types of telehealth uses (Tables 3 and 4). However, concerning the working-hours telehealth visits, and the after-hours visits—this decrease was found to remain significantly higher than the baseline level before the COVID-19 pandemic.

These findings support our assumptions that as expected, the use of all telehealth services was increased during the COVID-19 pandemic. The working-hours telehealth visits, which are the primary telemedicine meetings (with the personal physician), prospered during the COVID-19 pandemic, but even after this period, they remained at a higher level than they had before the COVID-19 pandemic period.

These findings confirm our hypothesis that older adults are interested in and capable of using telehealth services, given the opportunity and accessibility to do so—factors that were noticeably improved during the COVID-19 pandemic—despite their hesitations and barriers and the medical system’s mistaken belief that there is no point in teaching this population how to use digital health services.

In the univariate analysis, after classifying the number of telehealth services used into 2 categories (0 visits vs ≥1 visits), we found several factors that were significantly associated with increased telehealth use among the older individuals (Multimedia Appendices 1-3). Furthermore, we conducted a multivariate analysis, using a set of multivariable binary logistic regression models, which revealed that several parameters were significantly associated with the different types of telehealth utilization during each period (Tables 6-8).

Women were found to significantly use more working-hours telehealth visits with the personal physician, more after-hours telehealth visits, and more administrative telehealth services across all periods compared to men. This finding was supported by the multivariate analysis, showing that among men, there was decreased use of all types of telehealth services, during all periods. This finding may be explained by women’s greater general health literacy, eHealth literacy, and health awareness, as previously reported [24-26].

The findings for telehealth use by age groups in the univariate analysis were somewhat unexpected because it would be reasonable to presume that younger age groups (aged 65-74 years) would use telehealth the most. The findings indicate that the older age groups used telehealth services more frequently. Specifically, those aged 75 to 84 years used more working-hours telehealth visits during opening hours, more administrative telehealth use, and more after-hours visits following the COVID-19 pandemic. However, in the multivariate analysis, older age groups (aged ≥75 years) were found to use fewer working-hours telehealth services, fewer after-hours services, and fewer administrative services, during all periods.

These findings are consistent with the general belief that “younger, more highly educated and affluent seniors use technology more readily and across broader platforms than the older old, who as a group tend to be less affluent, less educated, and often have a significantly greater burden of chronic illness and disability,” as reported by Greenwald et al [21]. These authors claimed that younger seniors, who are more physically and psychologically integrated into the technological modern world, may have a more positive attitude toward the benefits of technology than older seniors. In contrast, the use of automated telephone menu systems, medical-related purchases on the internet (such as medical supplies or medications), and telemedicine videoconferencing with health care providers were all found to be more common among older adults than among younger adults in a 2011 study that looked into the type and frequency of technology use for a variety of health care activities [27].

The findings relating to SES in the univariate analysis were as expected: higher SES was significantly associated with increased telehealth use of all services investigated (working-hours telehealth visits, after hours, and administrative use) at all periods. This was also demonstrated in the multivariate analysis, in which people in medium and high SES significantly had increased use of all types of telehealth services during all periods. This could be ascribed to higher levels of education; increased knowledge and awareness of digital health services; increased income enabling the acquisition of computers, smartphones, and digital devices; and increased eHealth literacy linked to higher SES levels. These findings are in line with earlier studies, which found that people with lower median household incomes and less favorable insurance situations used web-based visits less frequently [26]. Another study also found a correlation between declining SES and declining probabilities of using telemedicine during elective surgery visits [28]. According to comparable findings among 16,000 patients with a new cancer diagnosis, those with the greatest SES were more likely to use telemedicine within 30 days of diagnosis [29].

The demographic sector findings in the univariate analysis were also a little surprising: the religious Jews, compared to other groups, had more working-hours telehealth visits with their physician, at all periods, and more after-hours visits (followed by the general Jewish population), before and during the COVID-19 pandemic. However, only in case of the administrative use, the Cherkess (Circassians) had greater use at all periods. These findings are unexpected given that using telehealth solutions in ultraorthodox communities offers several difficulties given their restricted access to virtual communication as well as their reluctance to engage in this novel modality of therapy. For many of these populations, receiving therapy through the web is strange and foreign, and it may be seen as going against their religious principles [30,31]. Nevertheless, in the multivariant analysis, those of the Jewish religion demonstrated highly increased use of all types of telehealth services during all periods.

The findings regarding the district that the older people belong to (where they reside) in the univariate analysis were inconsistent: before the COVID-19 pandemic, more working-hours telehealth visits were observed in the northern district; however, during and after the COVID-19 pandemic, more working-hours telehealth visits were made in the central areas. In addition, during all periods, more after-hours visits were made by residents of central regions. However, throughout
all periods, people from the southern district used telehealth more frequently for administrative functions.

Using the multivariate analysis, we saw that people living in the southern district used fewer working-hours telehealth services and fewer after-hours services during all periods but used more administrative services before and after the COVID-19 pandemic; people living in the northern district used more working-hours telehealth services before the COVID-19 pandemic but used less working-hours telehealth services during periods 2 and 3 and used fewer administrative services and fewer after-hours services in all periods; and people living in the central east district used less of all types of telehealth services during all periods. In general, we may conclude that a more peripheral place of residency was associated with reduced use of telehealth services. These findings are unexpected because one may anticipate that telehealth services would be used more frequently in remote places, where there are typically fewer health care personnel and resources.

The association of smoking status with telehealth use in the univariate analysis was interesting: for all types of telehealth services (working-hours telehealth visits with the personal physician, after-hours visits, and administrative telehealth use), people who had smoked in the past and stopped smoking (past smokers) substantially used more services compared to other groups (even more than nonusers) at all periods. This may be accounted for by such people’s greater health awareness, which led to a major change in their health-related behaviors. However, the multivariate analysis demonstrated that current or past smokers used more working-hours telehealth services, more after-hours services, and more administrative services during all periods. These findings could be additionally explained by the notion that current smokers have greater health problems, necessitating more frequent appointments with health care providers.

Finally, compared to persons without chronic diseases, people with chronic diseases significantly used digital services more frequently for administrative tasks, after-hours telehealth visits, and working-hours telehealth visits during business hours. This was true throughout all periods (including those before, during, and after the COVID-19 pandemic). This group’s rising use of telehealth was more than twice as high as that of people who were ordinarily healthy. These findings were also supported by the multivariate analysis demonstrating that the presence of any chronic disease was significantly associated with highly increased use of all types of telehealth services during all periods; in addition, the Charlson comorbidity score was found to be significantly associated with increased use of all types of telehealth services during all periods. This is not unexpected given that individuals with chronic illnesses or comorbid conditions may require additional help from health care professionals, and they frequently experience accessibility issues, making telehealth services particularly desirable to them.

In this study, we investigated the use of telehealth by older people who are still living at home to communicate with their clinic (administrative requests); individual physicians; or other medical professionals on evenings, weekend days, and holidays when clinics are closed. Nevertheless, there are telecare options available globally that enable the monitoring of patients with a chronic illness, such as heart failure, hypertension, diabetes, asthma, chronic obstructive pulmonary disease, or stroke [32-34]. In addition, more health care systems are using telemedicine video communication as a tool for health maintenance after discharge to lower hospital readmissions as well as expedited consult services (stroke, trauma, mental health screening, and surgical second opinions) [21]. Evaluation of telehealth programs for individuals, particularly the older adults, with chronic medical issues has yielded conflicting outcomes. Glycemic management and the use of health care services both benefitted patients with diabetes [35].

The research on telemedicine-enhanced emergency care for older people has been concentrated on residents of older people living community centers and has shown that high-intensity telemedicine services for acute illnesses have been effectively carried out, believed to be acceptable by older patients, and offered definitive care without needing a referral to the emergency department or urgent care [36]. An analysis of the impact of home-based telehealth interventions on the use of secondary care and mortality in a cohort of patients with COPD (chronic obstructive pulmonary disease), diabetes, and heart failure, most of whom (70%) were aged >65 years, led to lower emergency admission rates and lower mortality. [37] However, a different analysis of the same data revealed no impact on psychological outcomes or quality of life over a 12-month period [38].

Benefits of Telemedicine

It is well agreed that telemedicine-based care offers many benefits and advantages for older people. Older adults who use eHealth services can maintain their freedom and continue to live in their own familiar homes, where they feel secure and at ease. Their sense of security and quality of life are improved by being aware that they are constantly being watched and monitored [39-41]. According to Chou et al [40], older people who frequently used their telecare program and had better social welfare and health status also embraced using technology and had a higher quality of life. Their findings also showed that older people who believed telecare could help them with their health issues and were prepared to use it had a higher opinion of their quality of life, particularly in terms of their social interactions and home environment.

When an older person has decreased mobility, easy and quick web-based communication with a health care facility or professional becomes particularly crucial. This reduces travel time, speeds up diagnosis, reduces the need for repeat diagnostic tests and clinical services, and allows for older adult triage that is appropriate [42]. By offering some medical services at the patient’s home, telemonitoring and telerehabilitation help to decrease the frequency of hospitalizations and shorten their length, and the patient who is chronically ill can benefit by reducing the number of follow-up visits required at the medical facility [42-44]. In addition, due to telecare systems using eHealth services, an older person who is housebound can sign up via the web for a physician’s appointment, choose to receive a reminder of a due date for a medical consultation, receive straightforward remote medical recommendations or test results...
By allowing patients to develop and select the tools they will use as well as how they will use them, telecare fosters increased involvement [47]. Patients can actively participate in their own care through the use of telecare systems rather than just receiving it as a passive benefit. They become partners on an equal footing with their physicians, capable of making choices for their health on their own while being cognizant of the repercussions [48,49].

Through telemedicine, proactive healthy behaviors are formed. During an emergency, telemonitoring can identify pathological signs and symptoms and abnormal test results earlier than during or before a typical physician’s visit or examination, enabling prompt preventive action [42]. Telemonitoring also has a substantial impact on education. Patients who are conscious of their health status frequently begin to educate themselves on their disease and how to self-manage it. They have a better understanding of their body and are more motivated to modify unhealthy behaviors and live healthier lifestyles [42,50].

Telecare lessens socioeconomic and regional inequities in access to care as well as the uneven distribution of care quality. With the help of telecare technologies, older people can easily connect with a variety of functionally and geographically dispersed health care professionals at times that are convenient for both the patient and the health care provider [42].

Physicians in varied practice settings can easily access evidence-based medicine and effective clinical decision-making tools, such as knowledgeable colleagues in tertiary care facilities. In addition, patients can get the right type of care close to where they live, which is crucial for older people with chronic illnesses or disabilities because it affects their quality of life and level of care [42,51]. According to Chae et al [52], telecare was successful in lowering the frequency of clinic visits and also increased patient satisfaction in a trial of home health services for older people.

**Barriers to Telecare Use**

Although telecare offers undeniable advantages, it is important to understand that it also has limits, some of which are due to the older people themselves. The use of new technologies is frequently resisted by older individuals. Although computers and the internet have become important tools, older adults experience more trouble using them than younger people do.

According to a study by the Nielsen Norman Group, users aged >65 years had a success rate of just 53% when completing a series of assigned tasks (such as finding information and making a web-based purchase), compared to a group of younger users who had a success rate of 78%. In addition, the older group made an average of 3.7 errors on each task given, as opposed to the younger group’s average of 0.6 errors [53]. Another issue is the decline in cognitive and motor function that comes with aging (eg, vision, hearing, short-term memory loss, and physical impairment), which makes it harder to adapt to a changing environment and assimilate new behavioral patterns [54].

The strong desire to interact directly and personally with the physician is another trait shared by older people. They typically prefer face-to-face interactions with health care providers, so telemedicine-based services delivered remotely are frequently not seen as relevant to them. In a study on older adults aged ≥60 years, Bujnowska-Fedak and Mastalerz-Migas [55] found that 61% of older adults stated a strong preference for direct contact with medical professionals as the main deterrent to contacting their family physician, specialist, or nurses via telephone or the web. Eliminating in-person care may give older persons the impression that they are engaging in less social interactions. Resistance to telemedicine in older population may be a result of their concern that the new technology will negatively impact their social and personal relationships [56].

The next barrier is money. Pensioners, typically those who are in need, worry about the high prices of buying computers or other electronics. Older persons are frequently reluctant to spend money on home health care monitoring systems, despite the health advantages and long-term cost reductions made available by telecare services [55]. In addition, for older persons, privacy and security are top personal concerns. They want more assurance that their private information is kept private from prying eyes. Better health care is not as important to them as feeling assured about the security of their medical information [56,57].

**Accelerating Growth of Care Based On Telemedicine**

Technology advancements have made it feasible to put into practice solutions that, up until recently, looked to be a long way off. Israel is regarded as a highly developed country with excellent infrastructure, a high degree of entrepreneurship and innovation, and widespread knowledge of telemedicine services. As part of the Digital Israel Project, the Ministry of Health declared that one of its objectives was to “bring about a leap in the health system that will enable it to become sustainable, advanced, innovative, renewed, and constantly improving, by optimally leveraging the information and communication technologies available to the entire Israeli population” [58]. Similar to other developed nations, Israel has seen a rise in telemedicine use because of the COVID-19 pandemic.

Telecare improves the quality of life of older citizens with chronic illnesses worldwide by providing them with new options for education, prevention, diagnosis, treatment, and rehabilitation. It equalizes possibilities for patients from urban and rural locations and lessens socioeconomic gaps in access to care. According to a prior study [45], 41% of older people had a favorable opinion of eHealth services and were willing to use them if and when given the chance. A considerable shift from passive monitoring to more active use of telecare technologies that enable and promote direct connection has occurred in well-developed countries in recent years. Patients now have more control over their own health and welfare because of the changing health care system [43]. They can make health-related decisions on their own and with knowledge of the repercussions, and they work as equal partners with their physicians [59,60].

Nevertheless, widespread acceptance by older persons themselves is a crucial component in the development of telecare...
systems for the older population [43,51]. Despite the growing popularity of computers and mobile devices among older people and improved computer literacy, many of them are still unaware of the opportunities that telemedicine presents. Training for the use of telecare appears to be quite vital, as does ongoing education of the older population in this area. The needs, abilities, and preferences of older people should now be taken into account while providing telemedicine-based care, with adjustments made over time as care requirements change. Older adults have a wide range of needs, which can alter with time. Therefore, it is essential to individualize and adapt telecare systems for a range of abilities of older people, addressing their changing care requirements in a flexible and adaptable manner, always considering their impaired motor, sensory, and cognitive function.

Before using telecare technologies, all older people must be familiar with their utilization and aware of their advantages. In addition, perceptions of older people and their caregivers about the usability of home telecare are a substantial predictor of compliance with telecare [51]. An equitable health system should understand that while many older persons are willing and able to learn how to use telemedicine, for some, such as those with dementia and social isolation, in-person visits are already challenging, and telemedicine may be impractical. Clinics and geriatric modes of treatment, such as home visits, are crucial for these individuals [59].

Telecare will soon become a crucial aspect of older people’s lives, enabling them to function independently in a comfortable living environment, if technologies are developed that are familiar, usable, appealing, affordable, and fit into lives and plans of older people. Further research is required to accentuate the importance of offering the older population telehealth alternatives that are both easily accessible and easy to use.

Limitations

Despite the fact that this study included a large set of data from people belonging to Israel’s largest health management organization (and one of the largest health care organizations in the world), it still represents trends in telehealth use among the older population in Israel and not necessarily in all parts of the world. Furthermore, there were several parameters that could have an additional impact on telehealth use (eg, level of education); however, we did not have access to these data.

Conclusions

It is generally acknowledged that telemedicine-based treatment for the older people has several advantages. However, telemedicine also has limitations and barriers, some of which are due to the older people themselves.

The key findings of our study demonstrate that, despite all the challenges and hurdles, the older population uses telehealth services when they need them. People use telehealth services for administrative purposes more frequently, but they also consult with their own physician via telephone and the web and sometimes even use after-hours virtual consultations. These services make it easier for individuals to get medical care without having to travel, wait, or risk infection. Older people can increase their consumption as necessary during times of pressing necessity, such as the COVID-19 pandemic, or if they are afflicted with a persistent illness.

The study also reveals that even after the COVID-19 pandemic, most uses remained higher than they were before, implying that this population can learn how to use digital health services effectively and that they should be given the opportunity to do so by creating suitable and straightforward telehealth solutions tailored for this population and enhancing their usability.

Acknowledgments

The authors would like to thank the Clalit Research Institute team for their invaluable assistance in data extraction.

Data Availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Authors’ Contributions

MH was involved in the conceptualization of the study, methodology, investigation, data curation, analysis, and writing the paper. RS was involved in data curation, methodology, formal analysis, software, and supervision.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Univariate analysis (working-hours visits)—2 categories.
[DOCX File, 21 KB - aging_v711e52317_app1.docx]

Multimedia Appendix 2
Univariate analysis (administrative telehealth use)—2 categories.
[DOCX File, 19 KB - aging_v711e52317_app2.docx]
Multimedia Appendix 3
Univariate analysis (after hours visits)—2 categories.
[DOCX File, 20 KB - aging_v71e52317_app3.docx ]

References


Abbreviations

COPD: chronic obstructive pulmonary disease
OR: odds ratio
SES: socioeconomic status
VDI: virtualization desktop infrastructure

©Motti Haimi, Ruslan Sergienko. Originally published in JMIR Aging (https://aging.jmir.org), 24.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Experiences of a Community-Based Digital Intervention Among Older People Living in a Low-Income Neighborhood: Qualitative Study

Si Yinn Lu1*, MPH; Sungwon Yoon2,3*, MPH, PhD; Wan Qi Yee4, BSocSci; Nerice Heng Wen Ngiam4,5,6, MBBS; Kennedy Yao Yi Ng4,6,7, MBBS, MMEd; Lian Leng Low3,4,8, MBBS, MMEd, MCI

Abstract

Background: Older adults worldwide experienced heightened risks of depression, anxiety, loneliness, and poor mental well-being during the COVID-19 pandemic. During this period, digital technology emerged as a means to mitigate social isolation and enhance social connectedness among older adults. However, older adults’ behaviors and attitudes toward the adoption and use of digital technology are heterogeneous and shaped by factors such as age, income, and education. Few empirical studies have examined how older adults experiencing social and economic disadvantages perceive the learning of digital tools.

Objective: This study aims to examine the motivations, experiences, and perceptions toward a community-based digital intervention among older adults residing in public rental flats in a low-income neighborhood. Specifically, we explored how their attitudes and behaviors toward learning the use of smartphones are shaped by their experiences related to age and socioeconomic challenges.

Methods: This study adopted a qualitative methodology. Between December 2020 and March 2021, we conducted semistructured in-depth interviews with 19 participants aged ≥60 years who had completed the community-based digital intervention. We asked participants questions about the challenges encountered amid the pandemic, their perceived benefits of and difficulties with smartphone use, and their experiences with participating in the intervention. All interviews were audio recorded and analyzed using a reflexive thematic approach.

Results: Although older learners stated varying levels of motivation to learn, most expressed ambivalence about the perceived utility and relevance of the smartphone to their current needs and priorities. While participants valued the social interaction with volunteers and the personalized learning model of the digital intervention, they also articulated barriers such as age-related cognitive and physical limitations and language and illiteracy that hindered their sustained use of these digital devices. Most importantly, the internalization of ageist stereotypes of being less worthy learners and the perception of smartphone use as being in the realm of the privileged other further reduced self-efficacy and interest in learning.

Conclusions: To improve learning and sustained use of smartphones for older adults with low income, it is essential to explore avenues that render digital tools pertinent to their daily lives, such as creating opportunities for social connections and relationship
building. Future studies should investigate the relationships between older adults’ social, economic, and health marginality and their ability to access digital technologies. We recommend that the design and implementation of digital interventions should prioritize catering to the needs and preferences of various segments of older adults, while working to bridge rather than perpetuate the digital divide.

(JMIR Aging 2024;7:e52292) doi:10.2196/52292

**KEYWORDS**
digital divide; digital learning; smartphones; social gerontology; older adults; COVID-19 pandemic; technology adoption

**Introduction**

Globally, the COVID-19 pandemic has engendered unprecedented challenges for older adults. In addition to confronting the risks of infections and death [1], prolonged social distancing measures have worsened the physical and mental health as well as the quality of life in the older adult population. During this period, heightened loneliness, social isolation, and anxieties were attributed to reduced social engagement and support, mistreatment, and misinformation [2-4]. Loneliness is associated with higher risks of depression, functional and cognitive decline, and all-cause mortality and morbidity [5-7]. Older adults with lower socioeconomic status, preexisting comorbidities, and poorer access to health care were more likely to experience mental health–related problems [8,9].

Digital technology has surfaced as an essential way for older adults to cope with restrictions and mitigate loneliness by fostering social connections [10,11]. Studies have demonstrated the potential for technology to reduce perceived isolation via improved social support, connectedness, and engagement in activities [12,13] as well as promote a sense of autonomy and confidence [14]. Older adults who had more frequent digital contact, for example, through video calls, social media, or phone calls with friends and family, during the pandemic reported higher levels of satisfaction and well-being and a lower sense of social isolation [7,15,16]. Digital interventions have also been developed to enhance the psychosocial well-being of older adults. For example, a digital human facilitator was found to be feasible and acceptable in alleviating older adults’ stress and loneliness [17], and a digitally guided group intervention increased their digital literacy and abilities to cope with distress [18].

However, a large proportion of older adults globally continue to have little or no access to digital technology [19,20]. With the accelerated digitization of basic services (eg, financial transactions, health, and communications), the growing digital divide may further worsen the inequities in health and exclude the already marginalized segments of the population. Inequalities in access to digital literacy exist not only based on age but also based on income, education, sex, disability status, and immigration status, which culminate in “distinct positionality of privilege and disadvantage,” shaping diverse attitudes and environmental contexts that shape these attitudes and perceptions toward digital technology.

To fill this gap, this study aimed to explore the experience of and perceptions toward a community-based digital intervention among older adults living in low-income neighborhoods during the COVID-19 pandemic. Particularly, we sought to understand the motivations, perceived challenges regarding the adoption and continued use of smartphones, and benefits derived from the learning process, all in the context of social and economic vulnerabilities. The findings will inform how older adults who are considered vulnerable can most benefit from digital technology adoption and how program implementers can tailor the design of the interventions to older adults’ needs to maximize their effect.

**Methods**

**Overview**

This qualitative study was part of a larger mixed methods study conducted to evaluate the impact of a community-based digital intervention, namely, Project Wire Up, on older adults’ digital literacy and health-related outcomes in low-income neighborhoods in Singapore. This study used a generic qualitative methodology because the aim was to explore how older adults interpret and attribute meaning to the community-based digital intervention and smartphones [29].

Aligned with a constructivist epistemology, this methodology...
was selected to emphasize and foreground older adults’ subjectivities and their experiences regarding the intervention.

**Ethical Considerations**

Ethics approval to conduct the study was obtained from the SingHealth Centralized Institutional Review Board (2020/2722). Ethical considerations were accounted for in the design and implementation of the study within the context of a marginalized older adult community. For example, as in-depth interviews with older adults regarding the challenges they face in their everyday lives may elicit sensitive or upsetting recalls, interviewers will provide options for receiving various forms of support when needed (eg, taking breaks, postponing the interview, and options for referral to mental health resources) and remind participants that responding to interview questions is voluntary. In addition, as participants are referred to the study upon completion of the digital intervention, they may feel obliged to enroll in the study. Thus, the study team will highlight that participation in the study is voluntary and will not influence their access to services or ability to participate in subsequent community interventions. Informed consent was obtained before the interviews.

**Setting**

Singapore is experiencing an accelerated rate of population aging [30]. By 2030, the country will have one of Asia’s oldest populations, with one-fourth of the population aged ≥65 years. During the COVID-19 pandemic, heightened safety measures were implemented from April 2020 to June 2020 to curb widespread transmission. Nonessential services were stopped, and movement in public spaces was significantly curtailed. As most of the populace turned to digital means to maintain social connections and keep up to date with the news, a digital divide based on age became palpable [31]. Concerns that older adults would be left behind in the digital world became an impetus for a slew of government programs such as the Seniors Go Digital [32] to encourage smartphone adoption among this group [33,34]. Considering the heightened digital exclusion experienced by older adults with socioeconomic disadvantages, efforts were dedicated to enhancing access to digital literacy for this group of older adults [35].

Project Wire Up was established in July 2020 by TriGen, a voluntary organization, in partnership with Singapore General Hospital, Infocomm Media Development Authority, and older adult activity centers in Singapore to pilot a volunteer-led, one-on-one, home-based digital literacy program [35]. Those enrolled in Project Wire Up were eligible to purchase smartphones at subsidized rates and were matched to volunteers who provided one-to-one coaching regarding digital skills, including making video calls, connecting to Wi-Fi, performing web-based purchases, and using government services. An average of 6 sessions were conducted at participants’ homes over 3 months and were personalized according to their interests and competencies. Upon the completion of the program, volunteers may connect older adults to formal and informal networks for social support [35]. The aim of the intervention was to close the digital gap and improve social connectivity among older adults with lower socioeconomic status living in public rental neighborhoods. Public rental housing units, a sensitive indicator of area-level socioeconomic status in Singapore, are heavily subsidized flats that cater to lower-income households. Individuals eligible for public rental housing have total household gross monthly incomes that do not exceed approximately US $1000 [36].

Older residents in public rental flats face higher risks of frequent hospital admission and readmission, higher use of hospital and emergency department services, and longer durations of hospital stay [37]. Living in rental flats has also been correlated with poorer physical and mental health, including poorer cognitive function and higher depression rates among older adults [38]. In addition to shouldering a higher disease burden, they are more likely to have limited social and financial support. As poorer digital literacy may negatively influence older adults’ health via weaker social connections [39] and impact their ability to access health information and increasingly digitized modes of health care delivery [20], Project Wire Up’s primary goal of enhancing social connectivity and digital literacy has the potential to improve older adults’ well-being and mitigate health care use.

**Recruitment and Data Collection**

The study team worked closely with the implementation team to recruit eligible participants who met the inclusion criteria of being aged ≥60 years and successful completion of the program. Participants aged ≥60 years or who had not completed the intervention were excluded. A list of eligible participants was referred to the study team, who then contacted the participants to ask whether they were interested in participating in the study. A purposive sampling technique was used to obtain a diverse sample of participants, in terms of ethnicity, sex, and language spoken, to reflect the heterogeneous older adult population in Singapore. Between December 2020 and March 2021, the study team conducted semistructured interviews at various public rental housing estates. During the interview, participants were asked about their life histories, daily routines, challenges they faced during the lockdown, support received, meanings associated with smartphone use, and their experiences and challenges in participating in the program. Overall, 2 interviewers trained in qualitative research conducted the interviews, while observing the appropriate social distancing measures. Interviews usually lasted between 45 minutes and 1.5 hours and were conducted in English or local dialects (Mandarin, Cantonese, Hokkien, and Malay) and audio recorded.

**Data Analysis**

All interviews were transcribed and translated from local dialects to English. Then, the transcripts were coded using NVivo (version 12; Lumivero). Consistent with a generic qualitative methodology, a thematic analysis was conducted. An initial codebook was generated based on key sections of the interview guide. The interview guide drew on certain concepts from the existing literature, which posit that factors such as health, social network, and perceptions about technology influence how participants relate to and perceive smartphone use. These initial codes also included the challenges faced during the COVID-19 pandemic, as we anticipated that the pandemic may influence older adults’ attitudes and behavior toward technology. Participants were also asked about their experiences with the
program, including preprogram expectations, memorable moments, preferred mode of learning, postprogram smartphone use, and what they hoped to learn in future. These program-specific questions were intended to inform the design and implementation of future interventions.

Using this initial codebook, 2 coders coded 3 transcripts together through inductive and inductive coding methods to identify regularities in ideas and other emergent themes that may be relevant to the research problem. In this process, existing categories were expanded, and new categories were added to the codebook. Then, both coders coded the transcripts separately using the reviewed codebook. As both coders have each conducted several interviews with study participants, they had an in-depth understanding of and familiarity with the data and conducted regular discussions after coding 4 to 5 transcripts to resolve any discrepancies and discuss whether the key analytic categories that emerged were reflective of the meanings expressed by the participants. Analytic memos were also written alongside the coding process to reflect on any issues that arose during the coding process, emergent patterns, and thematic categories and subcategories [40]. Constant comparative analysis was performed throughout the process to compare the interview data to emerging categories and to determine the consistency in coding. Categories were created and refined when the data did not fit into categories. Then, the coders, together with members of the study team, discussed the key themes that were most salient in the interview data that shaped older adults’ experiences with the intervention and smartphone use. This paper followed the COREQ (Consolidated Criteria for Reporting Qualitative Research) checklist to ensure comprehensive and transparent reporting of the results [41].

Results

Participant Characteristics

We conducted a total of 19 interviews over a period of 4 months. Data saturation was achieved at 15 interviews, and we conducted a few more interviews to ensure that new data did not disclose new insights. Of the 19 participants, 12 (63%) were female participants, 16 (84%) were Chinese, 13 (68%) were widowed or single, and 15 (79%) lived alone in 1-room rental flats. Most participants reported a relatively low educational attainment, where 32% (6/19) of the participants received no formal education, and 26% (5/19) received primary education. The participants’ characteristics are presented in Table 1.

In this study, we identified three themes that reflect the perceived challenges and benefits related to learning the use of smartphones among older adults who live with low income: (1) age and social marginality, (2) technological design as a form of exclusion, and (3) digital learning process as a tool for mitigation of social isolation and marginality.
Table 1. Participants’ characteristics (N=19).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Participants, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
</tr>
<tr>
<td>60-69</td>
<td>5 (26)</td>
</tr>
<tr>
<td>70-79</td>
<td>7 (37)</td>
</tr>
<tr>
<td>≥80</td>
<td>7 (37)</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>7 (37)</td>
</tr>
<tr>
<td>Female</td>
<td>12 (63)</td>
</tr>
<tr>
<td><strong>Ethnicity</strong></td>
<td></td>
</tr>
<tr>
<td>Chinese</td>
<td>16 (84)</td>
</tr>
<tr>
<td>Malay</td>
<td>1 (5)</td>
</tr>
<tr>
<td>Indian</td>
<td>2 (11)</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>4 (21)</td>
</tr>
<tr>
<td>Widowed</td>
<td>6 (32)</td>
</tr>
<tr>
<td>Divorced</td>
<td>2 (11)</td>
</tr>
<tr>
<td>Never married or single</td>
<td>7 (37)</td>
</tr>
<tr>
<td><strong>Highest level of education</strong></td>
<td></td>
</tr>
<tr>
<td>No formal education</td>
<td>6 (32)</td>
</tr>
<tr>
<td>Primary</td>
<td>5 (26)</td>
</tr>
<tr>
<td>Secondary</td>
<td>5 (26)</td>
</tr>
<tr>
<td>Vocational or diploma</td>
<td>1 (5)</td>
</tr>
<tr>
<td>University and above</td>
<td>1 (5)</td>
</tr>
<tr>
<td>No response</td>
<td>1 (5)</td>
</tr>
<tr>
<td><strong>Housing type</strong></td>
<td></td>
</tr>
<tr>
<td>1 room</td>
<td>15 (79)</td>
</tr>
<tr>
<td>2 rooms</td>
<td>3 (16)</td>
</tr>
<tr>
<td>3 or 4 rooms</td>
<td>1 (5)</td>
</tr>
<tr>
<td><strong>Living arrangement</strong></td>
<td></td>
</tr>
<tr>
<td>Staying with spouse only</td>
<td>3 (16)</td>
</tr>
<tr>
<td>Staying alone</td>
<td>15 (79)</td>
</tr>
<tr>
<td>Staying with helper</td>
<td>1 (5)</td>
</tr>
<tr>
<td><strong>Employment status</strong></td>
<td></td>
</tr>
<tr>
<td>Working (full time)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>Retired</td>
<td>13 (68)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>4 (21)</td>
</tr>
</tbody>
</table>

**Age and Social Marginality**

**Overview**

Older adults’ perceptions about smartphones and the digital intervention must be understood within the context of their experiences with aging and social precarities. Precarity refers to an existence characterized by insecurity, unpredictability, and vulnerability that could emerge at an intersection of social disadvantages that extend into later life [42]. This means that older adults’ everyday experiences of aging must be understood and situated within the social structures that they are embedded in. In this study, most participants were in the middle-old category and experienced social and income-related vulnerabilities. Thus, our analysis demonstrated that the manner in which older adults relate to smartphone use and define their
capacity as learners is inextricably connected to their age-related anxieties and perceived socioeconomic positioning.

**Aging-Related Precarities and Internalized Ageism**

The most cited barrier to learning was age-related cognitive and physical decline, including visual and hearing impairment, memory difficulties, and decreased agility in their fingers, which impeded the ability to use the smartphone and retain knowledge from each session. More than half of the participants felt embarrassed about not being able to remember what was taught. For example, some participants described the anxieties and feelings of helplessness related to a perceived “deterioration” in their minds that made it difficult to absorb and apply the information learnt—a problem they attributed to old age:

> You are young, your minds are good, you can put many things inside but our mind is deteriorating, old already. [I] cannot see, cannot walk fast, this is natural for every person who gets old. [Participant #3, female]

In addition to the challenges related to cognitive and physical constraints, participants also expressed negative aging self-perceptions throughout the interview. While these were not explicitly stated as barriers to smartphone learning, they reflect the attitudes toward learning or negative associations between learning and old age. For example, a participant expressed the futility and meaninglessness of learning during old age:

> At first, I did not want to take [the phone], [but] the manager at the SAC said, “take it, learn slowly.” I said I am already going to die, no point learning, so old already, no one will know what will happen tomorrow, just live one day at a time. [Participant #6, female]

Our participants considered their unsuccessful attempts at mastering smartphone use as a sign of their inaptitude due to old age and the futility of the learning process. Participants had a strong tendency to individualize responsibility for the outcome of smartphone learning. Despite feeling uncertain about their ability to sustain smartphone use, participants chose not to seek help from volunteers after each visit due to their fear of “troubling” them and their reluctance to be seen as a “burden.” For example, some participants assessed themselves as learners with no scope for growth and whose self-judgments exacerbate their poor self-conception, low self-worth (ie, as learners with no scope for growth and whose feelings of helplessness related to a perceived “deterioration” in their minds that made it difficult to absorb and apply the information learnt—a problem they attributed to old age:

> You are young, your minds are good, you can put many things inside but our mind is deteriorating, old already. [I] cannot see, cannot walk fast, this is natural for every person who gets old. [Participant #3, female]

Social Precarities and Self-Imposed Stigma

Participants’ perceived lack of interest and confidence to learn smartphone use were also shaped by the awareness of their social positioning in relation to other older adults. Overall, one-third of the older adults emphasized that it was particularly difficult for them to learn to use the smartphone because of language barriers and illiteracy. For example, 1 participant was quick to distinguish himself from those who spoke English and were literate—characteristics that he felt predisposed them to increased competency and ability to acquire smartphone skills at a faster pace:

> Those [literate] people know words, know English, know the language. It is different, teaching them is very fast. For us, we do not recognize words, you teach me 10 times but I cannot remember...Waste time, waste effort. [Participant #1, male]

Most Chinese older adults residing in the neighborhood where the intervention was conducted are Mandarin-speaking or dialect-speaking individuals, making it challenging for them to navigate the smartphone. While volunteers helped these older adults change the default language setting to Mandarin, some participants were still not able to identify the characters due to their limited literacy. Participants viewed smartphone use as being in the realm of the privileged “other,” which does not align with their identities as “low income” or “uneducated.” For example, when asked how she feels about using the smartphone for purposes such as seeing the physician will change her life, 1 participant expressed that smartphones were not suitable for the “kind of people” living in her neighborhood:

> Some of us are uneducated, if educated, they have means or help and the ability to have a higher [paying] job. They will not be living in this area, you must understand what kind of area people live and what kind of people are living inside here, it’s not only when you think it’s good you can see on TV it’s good for old people. [Participant #13, female]

By perceiving that smartphone use is not applicable to older adults living in a certain “kind of area,” some participants “classified” themselves as failing to belong to the “in group” of digitally savvy older adults, thus reducing their self-esteem and motivation to learn.

Technological Design as a Form of Exclusion

Participants also expressed challenges related to smartphone technology, with its design primarily catering to the needs of a certain type of digitally literate individuals. Many participants described their interaction with the device as a stress-inducing process, pointing out that the sensitivity of the touch screen, small font sizes, multiple apps, and colorful esthetics made it confusing to navigate the interface. For example, 1 participant lamented that a combination of poor eyesight and stiff fingers resulted in her accidentally dialing the wrong numbers, incurring the wrath of family and friends:

> The smartphone is very sensitive, my finger accidentally touches it, then the other person’s phone will get it, so that is the trouble for an old lady...[my] eyesight is very poor, if our eyesight is
poor and the writing is so small, how you expect us to see... This phone is sensitive. That’s why I told you it’s not suitable for us old people. [Participant #13, female]

Thus, participants noted the incompatibility of smartphone’s functions with the needs of older adults and suggested that smartphone use will be helpful for those in the young-old group but felt that learning at an older age may not be useful. The failure of some smartphone designs to consider older adults’ age-related cognitive and physical limitations, needs, and preferences may also account for their sense of ambivalence toward smartphone use. For example, some participants expressed a disconnect between the prospects of learning a new, potentially disruptive technology and the fundamental priorities that they have at this point in their lives.

When asked what it means to be healthy, dominant themes expressed by participants included the ability to “walk,” ability to eat as they desired, having a “clear” mind and good eyesight, absence of ailments and difficulties, and ability to “live day by day doing the things [they] enjoy.” Therefore, the purportedly transformative potential of smartphones was incongruous with what participants valued or perceived as essential to their current life situation:

I want my life to be as simple as possible, do whatever I want to do... I just want to be happy, my mind has no space for other complicated things. At my age, I do not know when I am leaving this world, learn already also no point. [Participant #8, female]

We do not use these phone applications, only the youngsters use, there is a camera, take whose picture? A lot of these games, play for what...? Young people like all these funny things, we old people only need big font, big screen, simple. [Participant #3, female]

However, 11% (2/19) of the participants were motivated by the opportunity to learn new skills and expressed comfort and familiarity with navigating the smartphone. Their motivations were primarily shaped by active early-life work experience and the desire to keep in contact with closely connected family networks. Support from family members also facilitated their experience of uptake and sustained use. This aspect holds significance as it shows that despite cumulative disadvantages structured by lack of education, income, and employment opportunities, social support could influence their access, familiarity, and motivation to engage in digital learning.

Digital Learning Process as a Tool for Mitigation of Social Isolation and Marginality

While older adults in this study encountered challenges influenced by their experiences of aging, social precarities, and technological barriers, some participants expressed benefits related to the strengthening of relational ties with new friends or family members.

Many participants stated that their interactions with volunteers were one of the program’s most memorable moments. They appreciated having “someone to talk to” and liked that the volunteers were friendly, helpful, and approachable and provided personalized attention to addressing their queries. Thus, participants enjoyed the social exchanges with the volunteers, particularly if they could “chat and get along very well”:

I miss him [the volunteer], because when I sit here, I miss him sitting next to me, talking to me, teaching me what to do, he is a good person, really good, he’s working, he said no problem auntie you can call me if you have any problems. [Participant #15, female]

By learning how to make phone calls, send voice messages, and use mobile phone apps, participants mentioned improved social relationships, such as making new friends through the volunteers, and found it easier to communicate with others using the video call function:

It is more convenient to communicate with people, you can see them. Otherwise at that time, we did not know how to use the function, then it felt like we lost contact. You can only hear the voices but not see the people. [Participant #12, male]

Thus, given that most older adults in this study lived alone, designing an intervention that not only emphasized acquiring smartphone skills but also concurrently nurtured social bonds and connections appeared to foster participants’ initial acceptance and adoption of digital technology.

Discussion

Principal Findings

This study explored the perceptions and experience regarding a community-based digital intervention among older adults residing in a lower-income neighborhood. We sought to understand older adults’ motivations for learning, challenges to uptake and sustained use, and benefits derived from the digital engagement processes in the context of age-related and social disadvantages.

We found that the meanings older adults associated with learning at old age shaped their motivation and confidence in learning. Participants cited cognitive and physical limitations such as visual or hearing impairment and memory difficulties and a perceived sense of futility and meaningfulness of learning at old age as factors that limited their capacity and interest in digital learning. Many older adults expressed embarrassment and frustration regarding not being able to retain or apply what was taught and viewed these challenges as a natural outcome of old age. These negative self-perceptions are emblematic of the internalization of ageist structures and stereotypes that associate “being old” with being not technologically savvy. These findings corroborate studies that have documented how older adults’ identification with the negative connotations related to their age group may deter technology use [44,45]. These ageist stereotypes that depict older adults as “inflexible” or unable to “adapt to new ideas and to the use of technology” contribute to older adults’ feelings of low self-efficacy and discomfort and beliefs that efforts to learn will be unproductive or embarrassing [46].

Our study also reveals that these age-based anxieties intersect with older adults’ experiences of other forms of marginalization, including perceived stigma of being less educated and literate.
than other segments of the older adult population. The finding that older adults encounter language barriers aligns with those of a study in Singapore that described how the fear of information and communication technology among older adults in Singapore may be explained by their “limited command of English,” particularly among those with lower levels of education and socioeconomic status [26]. This study also found that older adults in this group expressed concerns related to the affordability of purchasing digital devices [26]. However, a novelty of our findings is that although the presence of subsidized smartphones facilitated smartphone uptake, older adults’ lingering ambivalence and aversion toward smartphone learning is also shaped by their consciousness and perception that smartphone use and technology adoption belongs to a privileged other from which they are excluded, and this influences their interest and motivation to learn. While the literature has recognized the importance of income and education in influencing older adults’ intention to use technology [28], our findings contribute to the existing literature by highlighting that older adults may internalize ageist attitudes and stigma resulting from their positionality within the social and economic structure. These self-perceptions hinder their desire to participate in and sustain smartphone learning. In addition, in the context of Singapore, public messages also showcase digital savviness and active social engagement as a marker of successful and healthy aging [33,47]. As studies have shown that engagement in digital practices corresponds to sociocultural conceptions of aging that is “active, engaged, independent [and] highly productive” [48], we posit that broader norms of successful aging may shape the identities, feelings of disempowerment, and negative self-perceptions of older adults with lower income and, in turn, affect their dispositions and perceptions toward smartphone learning.

In addition, our study demonstrates that older adults’ difficulties in navigating the smartphone’s user interface point to the absence of age-friendly features. These findings confirm those of earlier studies that suboptimal design features of digital devices, such as display screens that are challenging to navigate, small icons, and overall “low levels of graphic design adaptation” that fail to cater to the needs of older adults, may completely inhibit access [49]. However, our study adds to the literature by emphasizing older adults’ perceptions that the design and function of smartphones are incompatible with their existing priorities, needs, and visions of “healthy aging” and their feeling that it would be more relevant to the needs of young or middle-aged individuals or those who are digitally literate. While our study reiterates the need to examine the heterogeneity in technology use within the older adult population [50], we also emphasize that older adults’ perceived irrelevance and lack of usefulness of the smartphone in their everyday lives provokes a broader consideration of how existing digital tools and technologies often cater to the preferences and needs of a certain segment of the population, while excluding others.

Our study also highlighted that for older adults typically living alone in low-income neighborhoods, the community-based digital intervention had significant social meanings, creating new forms of social connection and relationships. In addition to improving digital literacy, most participants valued the interaction with the program volunteers. The importance of relationship building within the scope of digital learning is also underscored by the few participants who had demonstrated a strong motivation to learn. Participants in this group consistently practiced and used the smartphone to stay in contact with family and friends. This pattern is evident in studies that indicate the importance of familial and social support, particularly the efficacy of intergenerational approaches in the learning of digital skills, where older adults learn digital skills more readily from their grandchildren [50,51]. Overall, our findings reinforce how older adults’ aversion to digital technologies are shaped by factors such as age-related barriers and negative self-perceptions and a lack of user-friendly digital devices. However, our findings add to the existing literature by highlighting how individual-level factors are intertwined with and situated within the structural vulnerabilities that older adults confront such as age and income-related stigma and marginality. Thus, this study makes an important contribution to the existing theoretical models of technology use. Models such as the Senior Technology Acceptance Model have considered how technological use may differ in the context of older adults’ age-related physical, psychological, and social circumstances that predict their attitudes and behaviors toward digital technology [52]. The Senior Technology Acceptance Model explains that technology adoption is influenced by factors such as older adults’ self-reported health and cognitive ability, social relationships and life satisfaction, and levels of self-efficacy and anxiety toward gerontechnology [49,52]. However, our study’s findings supplement these frameworks by underscoring how the attitudes toward smartphone use among older adult populations that are considered vulnerable must be contextualized within intersecting age-related and income-related precarities that contribute to internalized ageist attitudes and social stigma among older learners, thereby shaping their self-perceptions, motivation, and identities. Existing models of technology use should consider the systemic ageism or exclusion that particular segments of older adults experience that may hinder technology adoption. The consideration of how older adults’ attitudes and behaviors toward digital technology are impacted by their experiences of other systemic disadvantages urges a shift from placing the onus of digital uptake on older adults toward bolstering the existing technological systems and social supports to improve digital connectedness.

**Recommendations for Future Interventions**

Based on our findings, 4 key strategies could be recommended to enhance the teaching methods and learning outcomes of older adults in low-income communities, as described in the following subsections.

**Understanding the Compatibility of Digital Technology With Older Adults’ Lived Experiences**

Implementers should consider whether the digital intervention and device are compatible with participants’ needs, preferences, and social circumstances. To comprehend the factors that influence participants’ motivations and attitudes toward learning a new technology, implementers could conduct a needs assessment to identify participants’ healthy aging goals, daily
routines, support network, and interests to plan how the smartphone could be relevant to their priorities. Although the current intervention tailors the program according to participants’ interests and abilities (where tier 1 involves learning “basic” phone functions such as video calls and tier 3 involves more “advanced” features such as web-based purchases or using government services), participants may not be able to derive meaning from learning these functions if they are not useful in their everyday lives. For example, if a participant wishes to increase their social interaction, the volunteer can teach them the video call function so that they can make a call to a family member or friend. Smartphone learning should not be seen as an end but as a means to fulfill objectives that are of importance to older adults.

In addition, smartphones may also not be perceived as a “resource” amidst financially precarious circumstances, where the urgent trade-offs in everyday priorities of living (eg, inability to pay for medical bills and uncertainties of welfare apps) mean that smartphone learning and its uncertain “rewards” cannot be prioritized alongside other competing demands on time and energy. Thus, volunteers and older adults should jointly examine ways in which they identify the perceived utility and relevance of the digital innovation in their current routines and life goals. Moreover, it is crucial to recognize that older adults do not necessarily share commonly held assumptions of smartphone as the indisputably “better,” “more convenient,” and “simpler” option; even if they do, they may also not perceive or understand these terms in the same ways. Thus, program implementers should consult older participants about what they value, the types of meanings they ascribe to the purported benefits of the smartphone, and the types of learning approaches that can best meet these needs. Expanding the discourses and meanings attributed to smartphones by different subgroups, particularly individuals considered marginalized, can promote intervention frameworks centered on equity and social justice, thus refuting “structures and systems designed by and for persons in more communities considered socially disadvantaged.”

**Strengths-Based Approach to Dismantle Ageist Stereotypes**

To address older adults’ negative self-perceptions about aging, digital interventions should take a directive approach to dismantle ageist stereotypes before cultivating more independent forms of learning [53]. Opportunities to contemplate age-related challenges should be built into the learning model, enabling older adults to confront their self-perceptions as being a “less worthy” or “incapable” learner. For example, techniques such as motivational interviewing can be adopted by volunteers to better understand older adults’ motivations to learn or resist smartphone learning. Rather than using a deficits-based approach (eg, what older adults do not know), motivational interviewing [54] seeks to affirm participants’ strengths, wisdom, and values and develops a plan toward change based on their own insights. This approach emphasizes the creation of a nonjudgmental, respectful, and compassionate space, where the older adults’ choice to learn or not learn the use of smartphone is not frowned upon or stigmatized. When older adults feel more empowered to learn, they can begin to explore the possibilities of smartphone use and refute the previously held conceptions that technology use conflicts with their identities (ie, not for “someone like them”). At the same time, the encouragement of older adults to learn should not involve coercion or guilt-tripping those who choose not to participate. Efforts must be made to assure older adults that a lack of participation will not deprive them of any other community services or assistance, to reduce the likelihood of older adults participating out of fear or obligation. Moreover, the responsibility should not be completely placed on older adults to actively engage in and keep up with digital practices, where resistance to learning becomes stigmatized or viewed as a burden or sign of “backwardness;” the consequence would be a subversion of the “discourse of empowerment” that digital technology seeks to promote [48].

**Strengthening Social Ties Through Technology**

Our findings suggest that smartphones can be adopted to fulfill relational purposes. Incorporating the cultivation of relational ties, in the form of family members, peers, or volunteers within a digital learning model, can serve to be a “catalytic intermediary” to motivate the use of digital technology [26] and sustain older adults’ interest in the program. In addition, digital technology can act as a medium through which older adults who live alone or lack a supportive social network can expand their social capital by “forming new social relationships or maintaining existing social ties” [26].

Digital technology can also serve as a medium through which older adults can acquire new skills or habits as part of a learning group. For example, a study has explored the application of gamification techniques to encourage older adults to improve digital skills through interactive games with a partner on a touch screen tablet. These games were designed to improve cognitive and motor skills and facilitate social interaction and were found to be effective in improving the acquisition of digital expertise [55]. Thus, intervention models that incorporate problem-solving activities and collaborative peer learning can create an interactive space that nurtures social connections and diminishes feelings of loneliness among isolated older adults in communities considered socially disadvantaged.

**Ensuring Program Continuity**

We found that the lack of opportunities for continued practice and application reduced older adults’ motivation for sustained use of the smartphone following the intervention. Future interventions can provide options to participants based on their levels of interest, skills, and aptitude; this could include connecting older adults to guided learning groups to practice the skills taught or to specific interest groups (eg, playing mahjong on the web). In addition, volunteers can visit the older adults at a fixed time to resolve technical issues that they may have related to phone use.

**Strengths and Limitations**

While the literature has documented the challenges faced by older adults in the realm of digital learning, this is the first study that uses a qualitative approach to examine how older adults residing in low-income neighborhoods experience aging and the social and health-related challenges that facilitate or limit their self-efficacy and interest in digital devices. In a global
context, there have also been other types of interventions that focus on improving older adults’ digital literacy. For example, in North America, digital literacy training sessions have been conducted in public libraries and community organizations [45]; a 4-month program of weekly computer classes was organized for African American older adults with low income at an older adult community center to gain familiarity with assessing web-based information and privacy issues [56]; and a 4-week digital literacy program was conducted to equip older adults with knowledge about how to navigate their computer (eg, sending emails) during the COVID-19 pandemic [57]. However, to the best of our knowledge, it appears that no study specifically explored the impact of a home-based digital literacy intervention during the COVID-19 pandemic, particularly among older adults with lower income.

One limitation is that we were unable to analyze the data in terms of understanding how these experiences and perceptions regarding the program might have differed across different sociodemographic characteristics—sex, age profiles (young-old and old-old), and health conditions—which could have provided richer insights into the experiences of these subgroups. The distinct experiences of these subgroups and the types of responses needed to address the challenges they face also warrant further research [21]. While our study only considered the context of Singapore, we believe that these findings regarding the role of age and social and material precarities in shaping technological use and the suggested solutions to bridge the digital divide will be theoretically useful in understanding the experiences and perceptions of digital tools among marginalized populations in other contexts.

Studies should be conducted to develop culturally sensitive approaches that can promote digital devices as a potential resource that is relevant to the needs of deprived communities, for example, in ways that can potentially improve the socioemotional and physical health outcomes of individuals or serve as a coping strategy in a precarious environment. Mixed methods studies using implementation science approaches [45] should also assess the maintenance of digital interventions in low-income communities, understand what is suitable for whom, and devise educational frameworks specific to the teaching of digital skills that can empower older learners.

Conclusions

The findings illuminate the need for community-based digital interventions to be designed with the particularities of the older adults’ lived environment and experiences in mind and the sensitivity that these digital tools only occupy one facet of participants’ lives, alongside other priorities and needs. Further studies are required to understand how these dimensions can be integrated into the intervention to enhance the smartphone’s perceived relevance and utility, without being an unwelcome disruption. Measures aimed at promoting individual-level adoption of smartphones must also be addressed alongside approaches that tackle structural inequities, ageist structures, and stigma that disadvantages one group of older adults relative to others. Regarding those who choose not to participate in the “digital wave,” the society must be willing to find and support alternative solutions to include these older adults in ways that promote social contact, autonomy, and socioemotional well-being—outcomes that technology purports to achieve—while not perpetuating their exclusion.

Acknowledgments

The authors express their gratitude to all participants who provided insights for the study as well as Sharon and Dhiya for their assistance with data collection and analysis.

Conflicts of Interest

None declared.

References


https://aging.jmir.org/2024/1/e52292

--

Lu et al

JMIR Aging 2024 | vol. 7 | e52292 | p.438
(page number not for citation purposes)
2018; July 15-20, 2018; Las Vegas, NV. [doi: 10.1007/978-3-319-92034-4_43]

52. Chen K, Lou VW. Measuring senior technology acceptance: development of a brief, 14-item scale. Innov Aging

53. Maulod A, Lu SY. “I’m slowly ageing but I still have my value”: challenging ageism and empowering older persons through

54. Resnicow K, McMaster F. Motivational interviewing: moving from why to how with autonomy support. Int J Behav Nutr

55. Blažič BJ, Blažič AJ. Overcoming the digital divide with a modern approach to learning digital skills for the elderly adults.


57. Elbaz S, Gruber J, Elberhouni K, Bukhari SN, Rej S, Sekhon H. Evaluation of a virtual 4-week digital literacy program
10.1080/03601277.2023.2268499]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research

©Si Yinn Lu, Sungwon Yoon, Wan Qi Yee, Nerice Heng Wen Ngiam, Kennedy Yao Yi Ng, Lian Leng Low. Originally published in JMIR Aging (https://aging.jmir.org), 25.04.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Digital Literacy Training for Low-Income Older Adults Through Undergraduate Community-Engaged Learning: Single-Group Pretest-Posttest Study

Lisa M Soederberg Miller1, PhD; Rachel A Callegari1, BA; Theresa Abah2, PhD; Helen Fann1, MS

1Human Ecology, University of California, Davis, Davis, CA, United States
2Department of Gerontology, Sacramento State University, Sacramento, CA, United States

Corresponding Author:
Lisa M Soederberg Miller, PhD
Human Ecology
University of California, Davis
One Shields Avenue
Hart Hall
Davis, CA, 95616
United States
Phone: 1 5307523955
Email: lmsmiller@ucdavis.edu

Abstract

Background: Digital technology is a social determinant of health that affects older people’s ability to engage in health maintenance and disease prevention activities; connect with family and friends; and, more generally, age in place. Unfortunately, disparities in technology adoption and use exist among older adults compared with other age groups and are even greater among low-income older adults.

Objective: In this study, we described the development and implementation of a digital literacy training program designed with the dual goals of training low-income older adults in the community and teaching students about aging using a community-engaged learning (CEL) approach.

Methods: The training program was embedded within a 10-week CEL course that paired undergraduates (N=27) with low-income older adults (n=18) for 8 weeks of digital literacy training. Older adults and students met weekly at the local senior center for the training. Students also met in the classroom weekly to learn about aging and how to use design thinking to train their older adult trainees. Both older adults and students completed pre- and posttraining surveys.

Results: Older adults demonstrated increased digital literacy skills and confidence in the use of digital technology. Loneliness did not change from pre to postassessment measurements; however, older adults showed improvements in their attitudes toward their own aging and expressed enthusiasm for the training program. Although students’ fear of older adults did not change, their comfort in working with older adults increased. Importantly, older adults and students expressed positive feelings about the trainee-trainer relationship that they formed during the training program.

Conclusions: A CEL approach that brings together students and low-income older adults in the community has a strong potential to reduce the digital divide experienced by underserved older adults. Additional work is needed to explore the efficacy and scalability of this approach in terms of older adults’ digital literacy as well as other potential benefits to both older and younger adults.

(JMIR Aging 2024;7:e51675) doi:10.2196/51675

KEYWORDS
community-engaged learning; digital divide; underserved older adults; digital literacy training; intergenerational programs
**Introduction**

**Background**

Digital technology is a social determinant of health that plays a significant role in older adults’ lives, including their ability to engage in health maintenance and prevention activities; connect with family and friends; and, more generally, age in place [1-3]. Recent evidence from the COVID-19 pandemic indicates that the digital divide contributed to health inequities among older individuals who were unable to benefit from technology to support health and well-being when in-person alternatives were unavailable or unsafe [4-7]. For example, without adequate online alternatives, older adults experienced greater health challenges and increased social isolation [8,9]. Conversely, older adults who had technology support during the pandemic fared better. For example, results from a qualitative study within a continuing care community indicated that technology mitigated social isolation and loneliness during the pandemic [10].

Disparities in technology adoption and use are particularly pronounced among older adults compared with other age groups [11-13] and are even greater among low-income older adults [3,14]. Some evidence suggests that internet use among low-income older adults may be as low as 17% [11] and that health-related technology use is significantly lower among racial and ethnic minority older adults as well as among low-income older adults [15]. Barriers that contribute to low rates of technology use among older adults include broadband availability, cost of broadband and devices, lack of awareness of potential technology benefits, low self-efficacy, and lack of training [16,17]. Some estimates indicate that only 25% of older residents in low-income housing have reliable internet access [18]. However, another study on low-income housing residents showed that although the housing communities in the study had access to broadband, few of the residents used the internet [19]. Thus, providing broadband and a digital device is insufficient. Digital literacy training is particularly important within this segment of the population to overcome barriers and promote sustained engagement [20].

**Digital Literacy Training**

There have been numerous approaches used to teach older adults how to use technology [16,21-23]. Regardless of the approach, researchers tend to agree that hands-on training over multiple weeks and tailored training programs are particularly important to meet the needs of older adults [24]. Tailoring can be achieved in a variety of ways—either informally (eg, drop-in or as-needed help) or as part of an in-home or a classroom-style program in the community [25-29].

Two theoretical models of technology acceptance and use among older adults are particularly well-suited to promoting digital literacy in this population [16,24,30]. First, the Senior Technology Acceptance and Adoption Model (STAM) includes ease-of-learning and ease-of-using technology constructs and argues that these are significant drivers of actual use [31]. Second, the Center for Research and Education on Aging and Technology Enhancement (CREATE) model of technology use in later life focuses on older adults’ characteristics, including demographics and psychographics (psychosocial background), cognitive, perceptual, and psychomotor abilities. The model also considers the relationship between older adults’ capabilities and the demands of the technology task being performed (eg, more challenging tasks undertaken will require greater capabilities) [16,32]. Thus, the CREATE model informs training by highlighting the importance of the fit between trainee and training activities as well as the rate at which training proceeds.

Drawing on these 2 models, we developed a conceptual model of digital literacy training, which is shown below in Figure 1. First, our conceptual model incorporates the STAM’s [31] close connection between ease of use on the one hand and actual use on the other. Last, the conceptual model places training at the center of adoption as being critical for low-income older adults who are far less likely to have prior experience with technology use and learning to use technology. The model provides a starting place for considering how training can be tailored to meet the needs of older adults with varying backgrounds and capabilities. Second, our model relies heavily on the CREATE model’s [32] emphasis on older adults’ characteristics (eg, demographics and cognitive and perceptual abilities) and the importance of the relationship between older adults’ capabilities and the demands of the task being performed.
Training programs targeting low-income older adults are far less common in the literature. Nevertheless, research suggests that factors that support older adults in general may be even more important for older adults from low-income backgrounds who may be reticent to seek training. Given the wide-ranging prior experience with technology and learning environments, a higher degree of tailoring and personalization may be critical to training those with lower-income and education backgrounds [20,24,33-35]. In addition to the insecurities surrounding the use of technology, low-income older adults may be more likely to have insecurities surrounding their ability to learn to use a computer, which in turn can impact enrollment into educational programs [33]. Low-income older adults may also lack awareness of the potential benefits of technology, which could also negatively impact motivation to learn to use technology and to persist in learning, especially when challenges arise [36,37]. For these reasons, programs that provide one-on-one training with highly individualized content and pacing may be a particularly important way to engage this population of older adults in training programs. Textbox 1 shows the training features, supported by the literature reviewed above, which may be particularly well-suited to low-income older adult populations.

Intergenerational Approach to Digital Literacy Training

Another approach with the potential to engage and motivate older adults in digital literacy training, particularly older adults from underserved communities, is to provide opportunities for intergenerational relationships within the training process [38,39]. Intergenerational programs involve a younger generation (e.g., college students) interacting with older adults who are typically from the same community. Opportunities in which different cohorts or age groups interact are becoming increasingly less frequent in society due to age segregation and changing family structures. In an age-segregated society, education, work, and leisure are apportioned to younger, middle, and older ages, respectively, restricting opportunities for older adults to spend time with younger individuals [40]. Intergenerational opportunities are further limited by changes in family structure, including fewer children and an increased likelihood of relocating to pursue a job or other opportunities [41]. Intergenerational programs often focus on fostering...
cooperation, interaction, and exchange between generations and can provide benefits, such as improved social connectedness and attitudes toward aging for both students and older adults [42-47].

Support for an intergenerational approach to digital literacy training comes from a study in which 38 pairs of high-school students and older adults met 2 or 3 times for 1.5 hours of training [21]. Although the study did not include digital literacy outcomes, results showed satisfaction for both the students and older adults. In another study, researchers explored 3 approaches to digital literacy training with students (undergraduate upperclassmen students in professional programs [eg, pharmacy]) and older adults [48]. With the first approach, students trained older adults in a 30- to 60-minute appointment at the senior center on an as-needed basis. The second approach matched students in a gerontology class with older adults from a local Osher Lifelong Learning Institute for a minimum of 6 hours of training at times and places determined by each trainer-trainee pair. The third approach consisted of as-needed, drop-in sessions held by students in 2- to 4-hour time blocks at the Osher Lifelong Learning Institute. Results showed that among older adult trainees who attended ≥5 sessions, trainees valued the intergenerational relationships developed over the training period as well as increased interest in working with technology. Moreover, younger adults showed improvement in attitudes toward aging, confidence in teaching older adults, and comfort working with older adults [48]. Similar to the matching program in Leedahl et al [48], Arthanat [49] paired undergraduate older adults in an occupational therapy degree program with older adults for 3 months of training (8 training sessions total) as part of a service learning project for an assistive technology course. Again, the training schedule was set by each pair. Before training, students attended a laboratory session on technology and aging and were encouraged to use Facebook as a forum to connect with other student trainers across the semester. Researchers reported increases in older adults’ frequency of technology use for multiple purposes, including those related to health and hobbies, and self-reported independence in a range of digital activities, reflecting improved digital skills. In general, the data show that intergenerational programs can effectively engage older adults in digital literacy training.

Programs that bring together students and members of the community provide a critical opportunity for community-engaged learning (CEL). CEL incorporates activities typically associated with internships and service learning into formalized learning within courses that consider social needs and social changes in the community. It also emphasizes the significance of building relationships with individuals in the community to bring about social change [50]. Young adults are often at the forefront of social change, with a heightened interest in the broader world, where they see themselves in society and how they can make a difference [51]. Intergenerational CEL provides an opportunity for younger individuals to broaden their awareness of social needs by exposing them to issues related to digital exclusion among underserved older adults. Textbox 2 describes the intergenerational training features supported by the literature.

**Textbox 2. Summary of intergenerational digital literacy training used in this study.**

<table>
<thead>
<tr>
<th>Intergenerational content and features</th>
</tr>
</thead>
<tbody>
<tr>
<td>A community-engaged learning framework was used that emphasizes providing students with the opportunities to develop academic skills, civic competencies, and ethical commitments while exploring community-based efforts to address social justice issues.</td>
</tr>
<tr>
<td>The training program was integrated into a structured course with set weekly meeting times for (1) student classroom learning and (2) training older adults; set times remove weekly scheduling burden and uncertainty regarding the training schedule.</td>
</tr>
<tr>
<td>Same trainer-trainee pairs worked together across the training program to enable the trainer to get to know the trainee to foster a trusting relationship that is conducive to frank and open discussion about training needs and pace.</td>
</tr>
<tr>
<td>Weekly student classroom time used to teach students about aging (agism; technology as a structural determinant of health with cumulative disadvantage perspective; age-related changes in perceptual, cognitive, and motoric/physiological capabilities) and other factors that may impact older adults’ acceptance, adoption, and sustained use of technology.</td>
</tr>
<tr>
<td>Student classroom time was also used to teach students about design thinking and how to apply design thinking to designing personalized and tailored training for their older adult trainee.</td>
</tr>
</tbody>
</table>

**Background Work**

Before this study commenced, we explored the logistics of pairing students and older adults in the community within an existing course on aging and technology use. Students (n=30) were partnered with older adults (n=17) in small groups (typically, 3 students and 1 older adult) over 4 to 5 weeks. Students were responsible for contacting older adults by telephone (using a free app such as Google Meet) and scheduling weekly 30- to 45-minute meetings. Telephone discussions focused on older adults’ current use of technology, attitudes toward technology, and what their preferences for training would be if they were to seek training. Students made notes each week to track what they learned about their partner’s technology acceptance. At the end of the quarter, we asked both students and older adults to answer open-ended questions about the challenges and rewards of their student-partner interactions. Overall, we learned that both the students and older adults highly valued their time together, and their relationships grew as they got to know each other. However, scheduling weekly appointments took a substantial amount of time, due to difficulties in identifying time slots, cancellations, and rescheduling.
This Study
This study examined a digital literacy training program integrated within an intergenerational CEL course to explore a formalized pathway connecting undergraduate education and underserved older adults in the community. The digital literacy training program was embedded within a formalized course structure in which students received college credits as they learned about aging and social justice while working one-on-one with low-income older adults in the community. The training program, designed to work within a course structure rather than in parallel or as an add-on (eg, with a minimal connection between class time and training time), offered 3 additional benefits for student trainers and older adult trainees. First, a set time and place removed uncertainty in scheduling from week to week. Second, weekly contact between students and the instructor in the classroom provided time to teach students about aging, social justice, and digital literacy training. Third, weekly contact between students and older adults at the training location (ie, local senior center) allowed for meaningful relationship building. Fourth, the presence of the instructor at the training site supported the students and older adults by facilitating communication and troubleshooting should any problems arise. Overall, this provided students and older adults with formalized and consistent support and, potentially, higher-quality education for the students and training for older adults. Finally, we offered the CEL course as a first-year seminar (with no prerequisites) in an attempt to attract students early in their academic studies. Students from any majors (including "undeclared" majors) were able to review the list of first-year seminars and sign up for those they wished to take.

In pairing students with low-income older adults in the community over an 8-week period, we hypothesized that relationships would develop organically as the intergenerational pairs worked together. The CEL course included features to support the process and the course used “design thinking” to encourage students to consider the older adult holistically on how technology fits within this individual’s life. Design thinking is a human-centered approach that places the “user” at the core when solving “wicked” or ill-defined problems, such as how to design a digital literacy training program that is well-suited to the trainee [52,53]. The course also taught students about aging (eg, older adult characteristics shown in Figure 1), stereotypes and biases related to aging, and how to train older adults using a design thinking approach, which focuses on understanding the end user and defining the task at hand (ie, designing an effective training program for their trainee).

Methods

Participants

Undergraduate Students

There were no prerequisites and no restrictions on who could sign up for the course. Among the undergraduate students (N=26) in the course, 48% (n=13) were female individuals, and 77% (n=21) were underclassmen (ie, freshman or sophomore), and they represented a wide range of majors, including data science, mathematics, sociology, and animal biology. 77% (n=21) were underclassmen (ie, freshman or sophomore), and they represented a wide range of majors, including data science, mathematics, sociology, and animal biology.

Older Adults

Older adults (N=23) were recruited with the support of a local nonprofit organization as well as low-income housing organizations and the local senior center. Inclusion criteria consisted of older adults who were aged ≥60 years; were eligible for a federal or state safety-net program (eg, Meals on Wheels, senior low-income housing, Medi-Cal, Cal-Fresh [SNAP]); were residents in Yolo County; and had little-to-no prior experience with computers. Funds provided by the County helped pay for recruitment and enrollment support as well as new laptop computers (which the participants were allowed to keep) together with 2-year internet subscriptions for the low-income older adults in the training program. Screening resulted in 5 of 23 individuals being excluded (1 for not meeting the age criterion and 4 for not meeting the low-income criterion). The final sample of older adults (n=18) was predominantly female (n=17, 95%) and non-Hispanic White (n=10, 55%), with the remainder being Asian (n=4, 22%) and Hispanic White participants (n=4, 22%). Older adults’ age ranged from 61 to 87 (mean 72, SD 7.81) years and had a mean of 17.61 (SD 5.21) on the Lubben Social Network Scale, which assesses social networks for family and friends with possible a range of 6 to 36 and clinical cutoff of ≤12 [54].

Ethical Considerations

The study was deemed exempt by UC Davis’s institutional review board. Older adults were read an information script before the pretest, informing that they could quit anytime and that their individual-level data would be deidentified.

Measures

Undergraduate Student Pretest-Posttest Measures

Psychological Growth

The Psychological Growth scale (8 items) from the Attitudes to Aging Questionnaire [55] was used to assess students’ attitudes toward aging. The scale includes items, such as, “It is a privilege to grow old,” “As people get older they are better able to cope with life,” and “There are many pleasant things about growing older.” Responses are made on a 5-point Likert scale (1=strongly disagree; 5=strongly agree) and are summed, with higher scores reflecting more positive attitudes toward aging.

Fear of Old People

The Fear of Old People subscale of the Anxiety about Aging Scale [56] was used to assess students’ attitudes surrounding intergenerational relations. The subscale includes 5 items, such as “I enjoy being around old people” and “I like to go visit my older relatives,” with a 5-point Likert scale (1=strongly disagree; 5=strongly agree). Items were summed, with higher scores indicating more anxiety toward aging.

Working With Older Adults

We included 3 items to assess students’ attitudes toward working with older adults as they trained older adults [48]. The items were “I am comfortable working with older adults,” “I am confident in teaching older adults how to use technology,” and “I am likely to volunteer in the field of senior services,” with
responses on a 5-point Likert scale (1=strongly disagree; 5=strongly agree).

**Undergraduate Student Posttest-Only Measures**

**Rank Order**

Students rank-ordered five aspects of the course from most important to least important: (1) learning about aging; (2) human-centered design and design thinking; (3) community engagement; (4) working with older adults; and (5) getting to know their trainee, specifically.

**CEL Values**

CEL addresses a specific community interest, problem, or public concern; includes working with and learning from a community partner; connects and integrates community-engaged experiences with educational content; and includes structured critical reflection. Students were given the following prompt: “Now that you’ve experienced CEL, we would like your perspective on its value. Please indicate how important the following CEL characteristics are to you,” followed by 12 items, such as “It makes me a better student in the long run” and “I want to contribute to the good of our society,” with responses on a scale of 1 to 3 (not very important to me, score=1; neutral, score=2; and very important to me, score=3).

**Older Adult Pretest-Posttest Measures**

**Technology Skills**

The Mobile Device Proficiency Questionnaire (MDPQ-16) was adapted to focus on “laptop” proficiency rather than mobile devices [57]. The questionnaire includes 8 scales, each assessing how easily the individual can perform digital tasks (n=18), such as navigating onscreen menus using the touchscreen, sending pictures by email, finding health information on the internet, and entering events and appointments into a calendar with responses of never tried (score=1), not at all easily (score=2), not very easily (score=3), somewhat easily (score=4), and very easily (score=5).

**Loneliness**

Loneliness was assessed using the 3-item Loneliness Scale [58], which asks how often individuals feel they lack companionship, are left out, and are isolated from others, on a scale ranging from hardly ever (score=1) to often (score=3). Scores are summed to create an overall assessment of loneliness. People who score 3 to 5 are considered “not lonely,” whereas people who score 6 to 9 are considered “lonely.”

**Attitudes Toward Own Aging**

Attitudes Toward Own Aging is a subscale of the Philadelphia Geriatric Center Morale Scale [59]. Respondents are asked to indicate whether they agree (score=1) or disagree (score=0), with 5 statements about aging such as “Things keep getting worse as I get older” (reverse scored) and “I have as much pep as I had last year.”

**Older Adults Posttest-Only Measures**

**Relationship Quality**

On a scale of 1 to 5 (1=strongly disagree; 5=strongly agree), older adults were asked to rate the following: “To what extent do you agree with the following statements about working with the students?”; “I enjoyed working with the student(s)”; “I feel that I formed a close relationship with the student(s)”; and “I will miss interacting with the student(s) now that the project has ended.”

**Social Connectedness During Training**

Older adults were asked how they feel about social connectedness using the following statement: “Did you feel more socially connected during the technology training program?” Responses were rated on a 3-point scale (1=no, 2=a little, and 3=a lot).

**Self-Efficacy and Confidence**

Three items on the extent to which participants agree with statements on computer skill improvement and confidence were as follows: “Your laptop skills have improved.” “You are more independent when using your laptop.” and “Your confidence in using laptop has improved.” Responses were rated on a 5-point scale (1=extremely unlikely to 7=extremely likely).

**Overall Evaluation**

For overall evaluation, older adults were asked the following questions: “To what extent do you feel the training program was a positive experience” and “To what extent do you feel the training program was a valuable experience” with responses rated as 1=not positive or valuable to 5=very positive or valuable; “How likely would you be to recommend this technology training program to others?” with responses rated as 1=extremely unlikely to 7=extremely likely.

**Procedure**

**Course**

Undergraduate students enrolled in the CEL course in September 2022. In the first 2 weeks of the course, students (in groups of 1 or 2) were paired with an older adult so that each intergenerational group could begin to work together starting in week 3 of the course (which was week 1 of the training program) and continue to work together for the entire 8-week training program. Students met in the classroom on Tuesdays (50-min class) and in the senior center on Thursdays (50-min session), where older adults joined them for training. Figure 2 illustrates the back-and-forth nature of the education (CEL course) and training program and the overlap between the ten 50-minute classroom meetings and eight 50-minute training sessions at a nearby senior center where the training was conducted.
Each week, students submitted 2 assignments. First, they submitted their reflections on how the training process is going, observations about needs for subsequent training, and what they learned about their trainee as a person. The last part (what they learned about their trainee) formed the basis for our design thinking approach, which placed heavy emphasis on understanding older adults, including where and with whom they live, family members, motivations and needs for technology, and any other information that enabled the student to understand (get to know) the older adult trainee. Second, students submitted a training record that documented each task and their step-by-step instructions. The training records were combined to create a learner booklet for the trainee at the end of the training program.

Students completed the pretest during the second week of the quarter and the posttest during the week 9 of the quarter via Qualtrics (Qualtrics International Inc) online software. For older adults, a research assistant administered the pretest via telephone during weeks 1 and 2 of the training program and the posttest 1 to 2 weeks following the training.

Digital Literacy Training
A nonprofit community partner, Yolo Healthy Aging Alliance, provided trainees with (1) funds to purchase broadband internet for 2 years and (2) a Chromebook (a relatively inexpensive laptop costing about US $150 each, with a larger keyboard and screen than most tablets), which trainees used during training and were allowed to keep after the training. At the first session, older adults were introduced to their student trainers and were given their new laptops. Students then helped the older adults establish a user account and practice using the keyboard and touchpad to navigate to various places (note: due to an error, computer mice were not delivered until after the third week). Starting with a list of tasks that were commonly mentioned in the preliminary study and in the literature (eg, how to send and receive emails), students worked with their older adult trainees in the subsequent sessions to determine an appropriate pace and topics of importance to the trainee. See Multimedia Appendix 1 for the overview provided to students. After each session, students created step-by-step instructions for each task that they covered that day, including reminders and screenshots with arrows to direct attention. Students shared these instructions with trainees at the following session to obtain their feedback on clarity and granularity (level of detail) and adjusted the subsequent instructions as needed to maximize the older adults’ learning. At the end of the training program, students presented these learner booklets to their trainees as a PDF document in hard copy and digital form (via email).

Results
Undergraduate Students
Pretest-Posttest Analyses
Table 1 shows the means for all the pretest-posttest measures. Students’ scores on the Psychological Growth scale of the Attitude on Aging Questionnaire were summed to create a summary score for each time point. We found that scores did not significantly change from the beginning (mean 29.58, SD 3.11) to the end of the course (mean 29.19, SD 3.86; t_{25} < 1, P = .46). Similarly, the sum of items on the Fear of Old People scale did not change significantly from the beginning (mean 10.19, SD 2.70) to the end of the course (mean 9.81, SD 2.95; t_{25} < 1, P = .42). Student ratings of comfort working with older adults significantly improved from pretest (mean 4.00, SD 0.57) to posttest measurements (mean 4.27, SD 0.45; t_{25} = −3.04, P = .006). Neither confidence in teaching older adults how to use technology (t_{25} = 1.31, P = .20) nor likelihood of volunteering in the field of senior services (t_{25} = 1.22, P = .23) significantly changed.
Table 1. Pretest-posttest summary variables for students (N=26).

<table>
<thead>
<tr>
<th>Pretest-posttest variables (range of possible scores)</th>
<th>Pretest measurements, mean (SD)</th>
<th>Posttest measurements, mean (SD)</th>
<th>t test (df=25)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Psychological growth (8-40)</td>
<td>29.58 (3.11)</td>
<td>29.19 (3.86)</td>
<td>0.73</td>
<td>.46</td>
</tr>
<tr>
<td>Fear of old people (5-25)</td>
<td>10.19 (2.70)</td>
<td>9.81 (2.95)</td>
<td>0.82</td>
<td>.42</td>
</tr>
<tr>
<td>Comfort working with older adults (1-5)</td>
<td>4.00 (0.57)</td>
<td>4.57 (0.54)</td>
<td>−3.04</td>
<td>.006</td>
</tr>
<tr>
<td>Confidence in teaching older adults (1-5)</td>
<td>4.15 (0.54)</td>
<td>3.96 (0.72)</td>
<td>1.31</td>
<td>.20</td>
</tr>
<tr>
<td>Likelihood of volunteering in the field of senior services (1-5)</td>
<td>3.38 (0.98)</td>
<td>3.58 (0.95)</td>
<td>−1.22</td>
<td>.23</td>
</tr>
</tbody>
</table>

Posttest Analyses

The end-of-course rankings of course features showed that students preferred working with older adults in the community over other course features: working with their older adult trainees was ranked first by 46% (12/26) of students and ranked first or second by 70% (18/26) of students. Working with older adults in general or specifically with their trainees was ranked first by 77% (20/26) of students. Learning about aging ranked the lowest with 77% (20/26) of students placing it in the bottom 2 positions. CEL rankings dropped in the middle (12/26, 46%, in the third position), and human-centered design rankings were evenly distributed across the 5 positions. Last, endorsements of CEL statements showed that students particularly valued contributing to the good of society, with 88% (23/26) indicating that this is very important to them (highest endorsement). The following 3 statements received very important ratings from 84% (22/26) of the students: “It helps me build compassion for myself and other people.” “The skills and knowledge that I gain will help me in my career,” and “I build relationships with people who live and think differently than I do.” In total, 80% (21/26) rated “It makes me a better student in the long run” as being very important. The following 2 statements received the highest number of not very important to me endorsements: “I learn from agents of change in my community” and “I believe it’s important to live out my faith.” Table 2 shows the scores for each value.

Table 2. Student ratings of community-engaged learning (CEL) values: not very important to me (rating=1); neutral (rating=2); and very important to me (rating=3).

<table>
<thead>
<tr>
<th>Statements</th>
<th>Scores, mean (SD); range</th>
</tr>
</thead>
<tbody>
<tr>
<td>I want to contribute to the good of our society.</td>
<td>2.89 (0.31); 2-3</td>
</tr>
<tr>
<td>The skills and knowledge that I gain will help me in my career.</td>
<td>2.85 (0.36); 2-3</td>
</tr>
<tr>
<td>I build relationships with people who live and think differently than I do.</td>
<td>2.85 (0.36); 2-3</td>
</tr>
<tr>
<td>It makes me a better student in the long run.</td>
<td>2.81 (0.39); 2-3</td>
</tr>
<tr>
<td>It helps me build compassion for myself and other people.</td>
<td>2.81 (0.47); 1-3</td>
</tr>
<tr>
<td>I can learn more outside the classroom.</td>
<td>2.74 (0.44); 2-3</td>
</tr>
<tr>
<td>My assumptions and beliefs are challenged, and I get to challenge others.</td>
<td>2.67 (0.54); 1-3</td>
</tr>
<tr>
<td>It empowers me to be an agent of change.</td>
<td>2.63 (0.48); 2-3</td>
</tr>
<tr>
<td>I see my community in new ways.</td>
<td>2.56 (0.50); 2-3</td>
</tr>
<tr>
<td>It informs the way I engage with the world.</td>
<td>2.56 (0.57); 1-3</td>
</tr>
<tr>
<td>I learn from agents of change in my community.</td>
<td>2.52 (0.50); 2-3</td>
</tr>
<tr>
<td>I believe it’s important to live out my faith.</td>
<td>2.44 (0.63); 1-3</td>
</tr>
</tbody>
</table>

Older Adults

Pretest-Posttest Analyses

The means of the summary scores are presented in Table 3. Overall, digital proficiency was analyzed in 2 ways. First, the MDPQ-16 scores were summed across the 16 digital tasks to indicate changes in overall proficiency. We found that scores changed significantly from pretest (mean 33.72, SD 14.05) to posttest measurements (mean 54.89, SD 14.42; t17=7.88, P<.001). Second, we assessed the changes in the range of activities performed by analyzing scores of 1 (never tried) on the MDPQ-16. Across all 6 activities, the number of “never tried” responses decreased from pretest (mean 9.61, SD 3.60) to posttest measurements (mean 3.83, SD 3.31), representing a significant change on a Wilcoxon signed-rank test of −3.73 (P<.001). Finally, we also examined the individual scales on the MDPQ-16 and found that all but basic skills (t17=1.48, P=.16) increased from pretest to posttest measurements, with P<.01 for the remaining 5 scales. Figure 3 shows the scores for all 6 scales at both assessment times.
Table 3. Pretest-posttest summary variables for older adults (n=18).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Pretest measurements, mean (SD)</th>
<th>Posttest measurements, mean (SD)</th>
<th>Test statistic&lt;sup&gt;a&lt;/sup&gt;</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDPQ-16&lt;sup&gt;b&lt;/sup&gt; total (16-80)</td>
<td>33.72 (14.05)</td>
<td>54.89 (14.42)</td>
<td>2.86</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Never tried activities (0-16)</td>
<td>9.61 (3.60)</td>
<td>3.83 (3.31)</td>
<td>−3.73</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Loneliness (3-9)</td>
<td>6.22 (2.58)</td>
<td>7.50 (2.77)</td>
<td>1.41</td>
<td>.18</td>
</tr>
<tr>
<td>Attitudes Toward Own Aging (0-5)</td>
<td>2.72 (1.74)</td>
<td>3.44 (1.50)</td>
<td>2.85</td>
<td>.01</td>
</tr>
</tbody>
</table>

<sup>a</sup>All tests are t test except “Never tried activities,” which was tested using the Wilcoxon signed-rank test.

<sup>b</sup>MDPQ-16: Mobile Device Proficiency Questionnaire-16.

Figure 3. Technology proficiency at pretest and posttest by Mobile Device Proficiency Questionnaire (MDPQ-16) scale.

Loneliness scores were summed across the 3 items (possible scores of 3 to 9), with higher scores indicating greater loneliness. As shown in Table 3, loneliness scores did not change significantly from pretest (mean 6.22, SD 2.58) to posttest measurements (mean 7.50, SD 2.77; t<sub>17</sub>=1.41, P=.18). Responses on the 5 Attitudes Toward Own Aging items were summed to create an overall score from 0 to 5, with higher scores reflecting more positive attitudes. We found significant improvements in attitudes toward aging from the pretest (mean 2.72, SD 1.74) to posttest measurements (mean 3.44, SD 1.50; t<sub>17</sub>=2.85, P=.01).

Posttest Analyses

Relationship quality ratings were positive, with 94% (17/18) indicating that older adults strongly agreed (highest possible score) that they enjoyed working with their student trainer, 89% (16/18) strongly agreed that they formed a close relationship with their student trainer, and similarly 89% (16/18) strongly agreed that they would miss interacting with their student trainer after the project ended (Multimedia Appendix 2). More than three-quarters of older adults (16/18, 89%) responded that they felt “a lot” more socially connected during the technology training program, with the remainder (2/18, 11%) indicating that they felt “a little” more connected. Older adults’ ratings of their computer abilities and confidence were also positive. Almost three-quarters of participants strongly agreed (highest endorsement) that their laptop skills had improved (13/18, 72%), that their confidence in using their laptop had improved (14/18, 78%), and that they were more independent when using their laptop (14/18, 78%). In terms of the overall evaluation of the training program, 100% (18/18) indicated their experience with the training program was very positive (highest endorsement) and 94% (17/18) indicated that the program was very valuable (highest endorsement). Finally, when asked how likely they would be to recommend the training program to others, 94% (17/18) responded extremely likely (highest endorsement) and 6% (1/18) responded very likely.
Discussion

Principal Findings
Multifactorial approaches are sorely needed to reduce the growing digital divide among low-income older adults [3,14,18]. The objective of the CEL study presented here was to address 1 component of the digital divide, digital literacy, using a potentially scalable approach. We sought to develop and implement a digital literacy training program using a CEL approach to bring together college students and low-income older adults in the community. There are several notable findings from this work.

First, we found significant improvements in digital literacy as assessed using the MDPQ-16, a measure of computer proficiency that is validated for older adults [57]. To our knowledge, proficiency has not been assessed in past work on training, regardless of whether the program used an intergenerational approach [23,48,49]. In addition to greater proficiency, we also found a significant increase in the breadth of technology use as reflected by a sharp decrease in the number of never-been-try activities after training. This finding is consistent with intergenerational and nonintergenerational training studies, showing an increased frequency of technology use across multiple tasks [23,49]. However, another study using a similar approach (ie, the item, use of technology in many ways) showed no change following intergenerational training [48]. Overall, our findings add to a small but growing literature showing that intergenerational technology training programs can be an effective approach to improving digital literacy [25] and add to this literature by extending the findings to include digital literacy benefits for low-income older adults.

Second, we found significant improvements in older adults’ confidence surrounding technology use, which is a critical component of technology acceptance and adoption [16]. This finding is consistent with past technology training studies showing improvements in older adults’ confidence [23] and enjoyment of working with technology [48]. Because we included assessments of both confidence and skill, this study extends prior research by showing that both can improve when using an intergenerational approach to promoting digital inclusion.

Third, our results showed beneficial effects on how older adults think about their own aging. Although past work has shown that intergenerational contact can promote positive attitudes among younger individuals [46], we are unaware of prior studies that examined these attitudes among older adults. Past research has shown that older adults’ positive attitudes toward their own aging protect against multiple diseases, including dementia [47,60]. Thus, policy makers interested in tackling the challenges of an aging population should consider investing in intergenerational programs to foster positive attitudes toward aging and enhance older adults’ well-being.

Fourth, findings from this study did not show significant improvements in loneliness from pretest to posttest measurements. Although this was somewhat surprising, other training studies have also shown no effect of training on older adults’ loneliness [23]. Indeed, a review of the literature on the effects of technology interventions, broadly defined, concluded that their impact on older adults’ loneliness is ambiguous [61]. One reason for our findings could be that the pretest survey was administered at the start of the program when older adults had already met their trainers and worked within the same room as other trainers and trainees at the senior center. By contrast, shorter training programs (eg, 8 weeks compared with several months) tend to show null effects on loneliness [61]. Although ours was a group program (in which participants met in a setting with several other individuals), it remains unclear whether loneliness levels would have been impacted by the training program. Therefore, future research is needed to explore the effects of group training programs on loneliness in older adults.

Finally, older adults’ end-of-program ratings suggest that the training program was a success. Self-reported improvements in digital literacy, program value, and program enjoyment were all high, and 94% (17/18) of older adults indicated that they were extremely likely to recommend the program to others. Enthusiasm for working with college student trainers has been reported in the past work [25,48]. High satisfaction with our digital literacy training programs may also be because the low-income older adults in our study (1) were able to keep the laptop and (2) received funds for a 2-year broadband subscription following the training (this support was for a federal low-income program called the Affordable Connectivity Program). This may also have helped to increase engagement and commitment to learning how to use the technology. Policy makers and community organizations interested in bridging the digital divide among low-income older populations should provide tangible support, such as digital devices and broadband connectivity to enhance program outcomes and promote continued digital engagement.

The significant improvements observed in older adults’ digital proficiency, confidence in technology use, and attitudes toward aging underscore the potential of intergenerational approaches that bring together older adults and college students to promote digital inclusion and well-being among older adults. Programs that help to formalize opportunities for undergraduates to work with low-income older adults as part of their undergraduate education (rather than in addition to it) may be particularly impactful for both older and younger adults. In subsequent sections, we outline some of these benefits.

CEL Approach

Reduces Uncertainties Surrounding Logistics of Meeting Times and Place
The prescheduled meeting times and a meeting place at a local senior center provided structure, which reduced uncertainties for both students and older adults. Data from our preliminary study indicated that both students and older adults were frustrated and confused by scheduling constraints and last-minute changes.

Train the Trainer
Another advantage to this approach is the ability to focus on training the trainers, which in this study included students learning about design thinking and aging. The ongoing
educational support to enable students to learn how to train older adults may be even more important when working with low-income older adults. Although other technology training programs have included educational support for students, they included only 1 session [48] or 1 session along with an optional social media forum for student trainers to support each other [49]. In this study, students’ questions and observations about how to tailor the training to be more effective continued throughout the duration of the program.

Focus on Social Good

As with public psychology [62], CEL is concerned with social good and the welfare of others. This importance of social good is becoming increasingly acknowledged in other public and private institutions of higher education [63]. In this study, we found that the students were also concerned with social good. CEL ratings reflect students’ interest in community, social welfare, and serving those in need. Developmentally, young adults tend to be concerned with the need for social change and identification of ways in which they can make a difference [51]. CEL with an intergenerational focus provides an opportunity for younger adults to express their concerns and broaden their awareness of social needs via exposure to underserved older adults. Furthermore, when CEL is embedded within the curriculum, students can more easily take advantage of opportunities to “give back” while also working toward their academic goals. This may be particularly important for underrepresented students who are often unable to do internships for a variety of reasons, including strict timelines for graduation, work obligations (sometimes several jobs), and family responsibilities.

Community and Campus Relationships

CEL relies heavily on close working ties with community members. In this study, the most important community members were the low-income older adults. Indeed, the relationship between the older adult trainees and the student trainers played a critical role in the success of the program. Community organizations also play a critical role and can help with recruiting older adults; donating space and associated services (eg, tables, chairs, parking, signage, and communication); and applying for funding to purchase laptops and broadband subscriptions. Finally, campus stakeholders also play a central role in the developing, testing, and scaling up of programs that focus on underserved older adults in the community. For example, campus leaders can express their enthusiasm for CEL courses, provide CEL experts to support instructors interested in designing courses, develop relationships with community stakeholders, and provide financial incentives (eg, to programs and departments) to incorporate these courses into existing degree requirements. The importance of campus and community partnerships cannot be underestimated [64]. Future work should therefore consider innovative ways to engage community organizations and campus leaders in efforts to build effective and sustainable intergenerational programs that improve low-income older adults’ digital literacy.

Intergenerational Approach

This study used an intergenerational approach to serving low-income older adults. As suggested by the posttest scores, students valued getting to know their older adult trainees. Their experiences across the program led to greater confidence in working with older adults, as has been shown in past work [48]. In addition, consistent with past research, we did not find significant changes in fear of older adults ($P=.42$) [44,48]. Surprisingly, however, we did not find significant improvement in students’ attitudes toward aging as has been reported in the past assessments [48]. A closer look at the prior study indicates that students in this study scored as high on the Psychological Growth scale at the pretest time point (mean 29.58, SD 3.11) as those in the study by Leedahl et al [48] at the posttest time point (mean 29.42, SD 3.19), suggesting that there was little room for improvement as a result of their interactions across the training program. In addition to whether changes in attitudes occur within an academic term, it is important to consider the possibility of long-term effects of intergenerational programs, including CEL. In this study, the first-year seminar course had no prerequisites, potentially attracting those who would not otherwise consider working with older adults. Given an increasingly age-segregated society [40] and a severe shortage of individuals trained to work in fields related to aging [65], this approach may promote interest among students to work on solving some of the pressing issues in facing a world with an unprecedented number of older adults [66,67].

Limitations

There were several limitations of our study. First, because of our single-group design, we cannot know which components of the training program were responsible for the beneficial outcomes. For example, we cannot disentangle the effects of training from the effects of owning a new laptop or assume the intergenerational component is superior to other models of training (eg, the use of technology experts). Second, our pretest measures might have shown a stronger impact on the program had we administered them 1 or 2 weeks before the start of the training program. Because we assessed them in the first 2 weeks, it is possible that some outcomes, such as loneliness and basic technology skills, might have already improved. Third, as with some previous training studies [25,48], the sample size was small, which can limit the ability to detect smaller effects. The sample was also predominantly consisted of White and female individuals, limiting the generalizability of the findings. Our sample does not, for example, allow us to account for the potential role that sociocultural factors (eg, culture, country of origin, intersecting identities, situations, or the interplay between these factors) play in the augmentation of digital literacy in older adults. Importantly, however, participants in this study were from low-income households, which are significantly underrepresented in the literature. Finally, our study did not include a follow-up to examine the long-term impact of training on new technology adoption and sustained use of the laptops. It is also possible, for example, that greater digital literacy skills would lead to decreased loneliness over time as older adults begin to use technology for social interactions [68]. Additional work is needed to explore the longer-term impact and scalability of this approach to promoting digital literacy among low-income older adults.
older adults and to examine other potential benefits to both older and younger adults.

Conclusions
Taken together, the current research contributes to a growing body of research on digital literacy training and provides a potential pathway to address the digital divide among underserved older adults [3,14,18]. Digital inclusion is central to older adults’ ability to remain independent and live in their own homes as they age. Moreover, the intergenerational CEL approach used in this study promotes mutual respect across generations, breaks down harmful stereotypes, and helps to build a stronger community. Moving forward, continued research in this area is crucial for informing policy decisions that support digital inclusion for older adults and help to address broader challenges related to an aging global population, digital fairness, social justice, and the shared fate of humanity.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Digital literacy training content and process overview that students used while training older adults.
[DOCX File, 54 KB - aging_v7i1e51675_app1.docx ]

Multimedia Appendix 2
Older adults’ posttest-only data.
[DOCX File, 78 KB - aging_v7i1e51675_app2.docx ]

References


52. Miller et alJMIR AGING


Abbreviations

**CEL:** community-engaged learning
**CREATE:** Center for Research and Education on Aging and Technology Enhancement
**MDPQ-16:** Mobile Device Proficiency Questionnaire
**STAM:** Senior Technology Acceptance and Adoption Model
Long-Term Adoption of Televisits in Nursing Homes During the COVID-19 Crisis and Following Up Into the Postpandemic Setting: Mixed Methods Study

Tobias Martin, MS; Sarah Veldeman, MD; Heidrun Großmann, PhD; Paul Fuchs-Frohnhofen, PhD; Michael Czaplik, MD, PhD; Andreas Follmann, MD

1Department of Anesthesiology, Faculty of Medicine, RWTH Aachen University, Aachen, Germany
2MA&T Sell & Partner GmbH, Würselen, Germany
3Docs In Clouds Telecare GmbH, Aachen, Germany

Corresponding Author:
Tobias Martin, MS
Department of Anesthesiology, Faculty of Medicine
RWTH Aachen University
Pauwelstraße 30
Aachen, 52074
Germany
Phone: 49 241 80 83137
Email: tmartin@ukaachen.de

Abstract

Background: There is growing evidence that telemedicine can improve the access to and quality of health care for nursing home residents. However, it is still unclear how to best manage and guide the implementation process to ensure long-term adoption, especially in the context of a decline in telemedicine use after the COVID-19 crisis.

Objective: This study aims to identify and address major challenges for the implementation of televisits among residents in a nursing home, their caring nurses, and their treating general practitioners (GPs). It also evaluated the impact of televisits on the nurses’ workload and their nursing practice.

Methods: A telemedical system with integrated medical devices was introduced in 2 nursing homes and their cooperating GP offices in rural Germany. The implementation process was closely monitored from the initial decision to introduce telemedicine in November 2019 to its long-term routine use until March 2023. Regular evaluation was based on a mixed methods approach combining rigorous qualitative approaches with quantitative measurements.

Results: In the first phase during the COVID-19 pandemic, both nursing homes achieved short-term adoption. In the postpandemic phase, an action-oriented approach made it possible to identify barriers and take control actions for long-term adoption. The implementation of asynchronous visits, strong leadership, and sustained training of the nurses were critical elements in achieving long-term implementation in 1 nursing home. The implementation led to enhanced clinical skills, higher professional recognition, and less psychological distress among the nursing staff. Televisits resulted in a modest increase in time demands for the nursing staff compared to organizing in-person home visits with the GPs.

Conclusions: Focusing on health care workflow and change management aspects depending on the individual setting is of utmost importance to achieve successful long-term implementation of telemedicine.

(JMIR Aging 2024;7:e55471) doi:10.2196/55471

KEYWORDS

telemedicine; televisits; telehealth; eHealth; electronic health; older adult care; nursing homes; change management; implementation science; technology transfer; innovation transfer; long-term adoption; COVID-19 crisis; postpandemic; coronavirus; digitalization
Introduction

Background

Population aging is an ongoing trend in industrialized countries and significantly impacts the health care sector [1-3]. Germany currently faces a steep increase in the number of older adult citizens and, concomitantly, in people needing care. Compared to the 2.63 million people needing care in 2013, this figure will increase by approximately 32% to estimated 3.48 million in 2030 [4]. As a consequence, more older adults will receive long-term care in nursing homes (NHs). To avoid future overcrowding of German hospitals and control overall health care costs, limiting hospital admissions (HAs) from NHs will be a major challenge. In fact, these HAs from NHs are frequently avoidable, with potentially avoidable HA cases from NHs accounting for €770 million (US $829 million) in avoidable health care costs in Germany each year [5].

Telemedicine is effective in reducing HAs from NHs and is an attractive modality of care, especially in the context of fewer home visits by general practitioners (GPs), a shortage of geriatricians, and difficulties in accessing health care in rural areas [6,7]. Both synchronous telemedicine, where a physician visits a patient in real time, and asynchronous telemedicine, where information on the patient is entered into a telemedical system and reviewed by a physician at a later time, can be implemented in NHs [8].

Another distinction of telemedical modalities is based on whether patients are accompanied by caregivers or not. Televisits are defined by the authors as videoconferencing among a remote physician, an NH resident, and an on-site caregiver (in this setting, a geriatric nurse) while having access to point-of-care (PoC) diagnostic devices that are integrated into the telemedical system. Televisits enable a structured physical examination and direct patient care, delegated by the remote physician and executed by the caregiver. This is not possible in video consultations, defined as simple videoconferencing between a physician and a patient. In the latter, the examination is limited without access to diagnostic devices, and direct patient care is not possible because there is no caregiver next to the patient.

Telemedicine is a safe modality of care and is noninferior to older adult patients presenting acute medical conditions in cases in which they are accompanied on-site by a nursing caregiver [9]. Several studies evaluating telemedicine for older adults have shown a reduced number of emergency department visits and HAs from NHs [10-22]. Despite these positive effects on the overall level of care and the widespread accessibility of telemedical tools, deployment in primary care and NHs is only progressing slowly.

Even though several implementation frameworks for telemedicine have been proposed in the literature [23,24], a lot of projects transferring telemedical or digital innovation into routine care fail due to poor consideration of change management (CM) aspects [25]. CM can be understood as an organized approach to drive organizational transformation from one current state to a new desired state. The concepts and various models commonly used for business transformation can be applied to health care, where new innovations are also constantly integrated [26-28]. While clinical research mainly focuses on creating evidence for better health-related outcomes, CM aims to ensure long-term adoption of change processes by promoting staff engagement and fostering a culture of continuous improvement. In fact, most of the clinical studies assessing the implementation of new technologies such as telemedicine focus on short-term adoption and technical issues but do not specifically address organizational, cultural, and educational challenges [29-31]. In contrast, the structured approach of CM involves methodically planning and monitoring the entire process to promptly identify and address challenges. This allows for the reduction in resistance to change, ensuring a smooth transition for health care providers and patients and, ultimately, enhancing the quality of care and organizational effectiveness. CM methods are action-oriented approaches focusing on managing specific change processes within one institution by directly addressing challenges. The field of implementation science (IS) must be distinguished from CM applied to health care. While IS also aims to understand implementation and the sustainability of implementation efforts, it takes a broader perspective. In fact, IS aims to create generalizable knowledge about effective implementation strategies of evidence-based health interventions by identifying facilitators and barriers across different contexts and health care settings [32].

Currently, many barriers to the implementation of telemedical tools are known, such as unstable internet connections and other technical issues, insufficient acceptance, privacy and security concerns, poor usability of the systems, a lack of patient support from health care professionals (HCPs), inadequate motivation and training, a shortage of staff, poor planning and engagement, and the fear of misdiagnosis or lack of trust in the technology [31,33-36]. However, implementation guidelines are still missing, and the best practice for implementing telemedicine for ensuring long-term adoption is still unknown. Moreover, there is a gap in research concerning the impact on nursing practice triggered by the organizational implementation of televisits.

Objectives

This study aimed to identify and address major challenges for long-term implementation of televisits as well as evaluating the impact of televisits on nurses’ workload and their nursing practice.

Methods

Setting

This study was conducted in 2 NHs located in 2 different rural areas of the federal state of North Rhine-Westphalia (Germany). Both NHs provide long-term stationary care for older adults, with an average resident age of > 85 years. Although they are comparable in size (NH 1 provides care for 93 residents, and NH 2 provides care for 90 residents), they differed in the number of employees within the study period (NH 1 had 64 nurses and nursing assistants for a total of 39.6 full-time equivalents, and NH 2 had 37 nurses and nursing assistants for 32.5 full-time equivalents).
equivalents) and in the number of trainee nurses (14 in NH 1 and 69 in NH 2). There were more women (86% of the employees in NH 1 vs 70% in NH 2, trainees excluded) and older employees (mean age 47.8 years in NH 1 vs 33.3 years in NH 2, trainees excluded) in NH 1 than in NH 2. NH 1 was built in the 1970s and is administered by a foundation managing NH 1, a local ambulatory care service, and 1 facility for assisted living for older adults. NH 2 is run by a local nonprofit organization with a history of >150 years that provides stationary care in 3 NHs. The organization that runs NH 2 also offers care in all the other relevant fields of older adults’ care (intensive care nursing, ambulatory care, assisted living communities, and palliative care). The regional district of NH 1 faces a steeper decline in the number of physicians than that of NH 2. This is reflected in the average travel time to the nearest GP practice, which was twice as long in the district of NH 1 as in the district of NH 2 (4.4 vs 2.0 minutes) in 2021 [37]. For the medical care of the residents, both NHs collaborate with a coordinating GP, who cares for approximately one-third to half of the total number of residents living in the NH. This is common in Germany, where the remaining residents are followed up by other GPs located within the catchment area of the NHs. Only the corresponding GPs, and not the GPs of the other 2 NHs, participated in the implementation of televisits. However, an NH resident followed up on by a GP other than the corresponding one could still benefit from a televisit if the coordinating GP was covering for other GPs on holiday or sick leave.

**Timeline**

The timeline of the covered implementation process during this study is shown in Figure 1. The total follow-up period between November 2019 and the end of March 2023 can be divided into 3 phases. The first phase between November 2019 and mid-August 2020 consisted in planning the implementation process, supplying and setting up the telemedical system, and initial training of the HCPs. This was followed by 2 distinct implementation phases: a short-term implementation phase until June 2021, during which restrictions due to the COVID-19 pandemic were still in place, and a subsequent long-term implementation phase after the COVID-19 restrictions had been lifted.

**Figure 1.** Timeline of the implementation process.

From the very beginning, an interdisciplinary steering committee (see the Evaluation: Steering Committee and Systematic Analysis of the First Implementation Phase section) organized and supervised the entire change process. After a requirement analysis, the telemedical system was first delivered in May 2020. The initial training was delivered via web-based seminars as COVID-19 restrictions did not allow access to the NHs for external persons. A total of 5 webinars lasting approximately 1 hour each and held consecutively every week covered all the relevant aspects of televisits (the topics were Introduction to telemedicine, Televisits, Televisits: the nurse’s perspective, Televisits: the GP’s perspective, and ECG, auscultation, and other diagnostic devices). In parallel, the battery lifetime was enhanced, a bigger screen was installed, and the internet connection in the NHs was improved. In August 2020, hands-on training was organized for the entire staff in both NHs. After a short theoretical introduction summarizing the main learning points of the webinars, the participants were trained in televisits using simulated scenarios. Workflow organization was not part of the initial training. The nurse managers in the NHs received advanced training to become superusers, enabling them to feel comfortable in performing televisits, administering the users in the software, and guiding other nurses in their learning process. They later acted as contact persons for their colleagues regarding televisits. After the hands-on training sessions in August 2020, the NHs were asked to perform—whenever possible—weekly televisits with their coordinating GPs while still maintaining the weekly on-site home visits. No specific instructions or guidelines were provided regarding how the televisits should be incorporated into the existing workflows of the NHs. This provided the NHs with the flexibility to integrate televisits according to their specific requirements and organizational contexts.

To allow for continuous training, the webinar sessions, as well as additional step-by-step guides, were made available on an e-learning platform. The nurse managers continued weekly
training of their colleagues and accompanied them in televisits when necessary. After a short familiarization period during which the nurse managers provided significant support, effective routine use was achieved between February 2021 and July 2021, covering the last COVID-19 lockdown in Germany from April 2021 to June 2021. When the COVID-19 restrictions ended, the first implementation phase was systematically analyzed within the steering committee. Challenges preventing further implementation were identified and addressed through specific control measures. This led to a new organization of televisits and marked the beginning of a second implementation phase beyond the constraints of the COVID-19 pandemic, referred to in this study as the long-term implementation phase.

The change process was monitored until the end of March 2023.

**Telemedical System**

Televisits were performed using the so-called TeleDoc Mobile system (Docs in Clouds TeleCare GmbH), a market-available and mobile medical cart system for televisits with integrated medical diagnostic devices. The TeleDoc Mobile system was equipped with a blood pressure meter (BU 540 connect; medisana GmbH), a blood glucose meter (MediTouch 2; medisana GmbH), a 1-channel electrocardiogram (ECG; WIWE pocket ECG; myWIWE Diagnosztika Kft), and an electronic stethoscope (Littmann stethoscope model 3200; 3M). A conference camera offered a 10-time optical zoom in high definition (PTZ Pro 2; Logitech International S. A.). Screens on both sides of the system allowed for synchronous video communication among the GP, the nurse, and the resident. The software version of the TeleDoc Mobile system underwent multiple updates during the study period from version 1.0 in 2019 to version 3.6.2 in 2023. Feedback from HCPs was directly incorporated into these updates to meet their specific requirements. Previous work on the TeleDoc Mobile system had already demonstrated the technical feasibility of televisits and good acceptance by users [38].

**Evaluation**

**Overview**

A mixed methods approach was used for evaluation at different time points, as shown in **Figure 2**. A project diary in paper format containing a questionnaire and televisit documentation protocols (Multimedia Appendix 1) served to assess the nurses’ baseline expectations and competencies and document the first televisits. This diary was distributed to all nurses in the NHs in May 2020, when the final telemedical system was delivered and the NHs were about to be trained and begin but had not yet done real televisits. The nursing staff were asked to fill in the questionnaire and document key elements of their first televisits, such as the date, the motive for consultation, and the number of PoC diagnostic devices used. The diaries were collected in January 2021. As already mentioned, the first implementation phase was observed and systematically analyzed within the steering committee in August 2021. At the conclusion of the study period in March 2023, qualitative interviews were conducted to retrospectively evaluate the nursing perspective regarding the implementation phases and the experienced change related to the implementation of televisits. The qualitative assessment of changes experienced in nursing practice was further quantitatively evaluated using a follow-up questionnaire in June 2023. The individual methods are described in detail in the following subsections.

---

**Figure 2.** Overview of the methods used for evaluation.
Questionnaire on Baseline Expectations and Competencies

The initial questionnaire (Multimedia Appendix 1) assessed nurses’ self-reported competencies in the use of both medical devices and computer programs. A scale from 1 to 6 following the German school grading system was used (1=very good, 2=good, 3=satisfactory, 4=sufficient, 5=poor, and 6=deficient). Baseline expectations regarding the impact of implementing televisits on the residents’ medical care and the nurses’ time spent on televisits were assessed using preformulated multiple-choice questions. The nurses were also asked to rate, according to the German school grading system and before starting real televisits, their self-perceived level of knowledge of the medical devices as well as their level of understanding of the TeleDoc software. As the questionnaire was linked with the televisit documentation protocols, the questionnaire was not collected directly after the nurses had received training but only with the diaries in January 2021. In total, 20 diaries (n=12, 60% in NH 1 and n=8, 40% in NH 2) were handed back for analysis. After 1 diary was excluded because of missing data, 19 (n=11, 58% in NH 1 and n=8, 42% in NH 2) were used for analysis. Single missing data points were imputed with median imputation. Only the results regarding the self-reported competencies and the initial expectations are presented in the main text of this paper; however, the data for all questions can be found in Multimedia Appendix 2. Continuous variables were described as medians and IQRs and compared using the Mann-Whitney U test. Categorical data were described as numbers and percentages and compared using the Fisher exact test. All statistical tests were 2-sided with an α significance level of .05. They were applied in Python (version 3.9.7; Python Software Foundation) using the scipy.stats package.

Telesvisit Documentation Protocols

Telesvisit documentation protocols (Multimedia Appendix 1) were included in the second part of the project diaries. The nurses recorded in these protocols when a telesvisit took place and whether the telesvisit was preplanned or organized for an acute issue, such as an acute medical presentation by a resident. The nurses also documented the reason for the consultation and indicated whether they would have contacted the emergency service or the GP out-of-hours service or waited until the GP was available if they had not had access to telesvisits. In a second part, the nurses were asked to indicate which medical devices they used and rate how well they got along with these devices and the software. Data presentation and statistical analysis were conducted in the same manner as for the data of the initial questionnaire (continuous variables: medians, IQRs, and Mann-Whitney U test; categorical variables: numbers, percentages, and Fisher exact test; 2-sided statistical tests; α=.05; Python version 3.9.7). The protocols ended with questions regarding the interaction with the GPs and how the residents felt about the medical care during the telesvisits. They also contained some open-ended questions for providing personal detailed responses. These data are not presented in this paper but can also be found in Multimedia Appendix 2.

Steering Committee and Systematic Analysis of the First Implementation Phase

An interdisciplinary steering committee monitored the entire implementation process. It consisted of members from clinical and change research fields (n=3 physicians doing clinical research and n=2 researchers in CM), a health insurance representative (AOK Rheinland/Hamburg, Düsseldorf, Germany), NH and nurse managers of the participating NHs (n=3), and technical developers from the telemedical system manufacturer (n=2). All authors were part of the steering committee. The steering committee conducted the requirement analysis and organized and developed all forms of training and information material for the nursing staff, such as the hands-on training, the e-learning classes, standard operating procedures for specific presentations, and pocket cards for the use of the medical devices. As the nurses in the NHs reported directly to the senior nurses of their NHs, they were indirectly represented in this committee. Monthly meetings were held throughout the follow-up period to discuss the ongoing process and challenges encountered. If the challenges could not be resolved immediately, specific committee members were assigned to elaborate possible control measures and action plans, which were then presented and discussed in subsequent meetings. The challenges and the resulting actions are presented descriptively. The figures describing this process were created using Inkscape (version 1.3, Free Software Foundation, Inc) and based on free images from Freepik [39] designed by the authors Freepik, Slidesgo (Freepik Company SL), and stories.

Qualitative Interviews

The COREQ ( Consolidated Criteria for Reporting Qualitative Research) guidelines were used to present the design and results of the qualitative interviews [40]. The COREQ checklist can be found in Multimedia Appendix 3 [40]. From March 2023 to June 2023, a semistructured interview study was conducted with nurses directly involved in the care of the residents (n=5 in NH 1 and n=3 in NH 2) and with senior nurses with additional administrative roles during the implementation of the telesvisits (n=1 in NH 1 and n=2 in NH 2). These senior nurses had all been part of the steering committee during the implementation process. The interview guide (Multimedia Appendix 4) was developed by TM, a physician and clinical researcher, and checked for consistency and missing questions by HG, a CM expert. The participants were selected through convenience sampling. The nurse managers proposed interview participation to all nurses who had performed telesvisits. The voluntary participants then received appointments for interviews via videoconferencing. The interviews took place in a dedicated room in the NHs where the nurses were alone and not interrupted during the interview. Before starting the interview, all interview partners were informed about the aim of the interview, which was to collect the individual perceptions and experiences related to the process of implementing telesvisits in their NH. All interviews were conducted by TM. While the senior nurses with additional administrative roles were known to TM before the interview, there was no previous relationship with the other interviewed nurses. The participants were also informed about TM’s background and that this study on telesvisit implementation was part of his research for gaining his degree as Doctor of
Medicine. They were also informed about his role in the steering committee, where he provided scientific guidance and assistance. The participants received information and gave informed consent for the interviews.

Most of the interview sections were dedicated to evaluating expectations, experiences, and perspectives regarding televisits as well as assessing the impact of their implementation on the nursing practice. These inquiries were all presented as open-ended questions. Another part focused on the addressed challenges and control measures adopted throughout the implementation process. To facilitate focused and structured responses, the nurses were asked to comment freely on predefined statements regarding various aspects of the implementation process within the study period.

The interviews were conducted once without follow-up or subsequent interviews. All the interviews were held in German, visually recorded, transcribed, and analyzed using thematic analysis. No field notes were made during the interviews or when reviewing the interviews. The transcripts were not returned to the participants for comments. As most of the questions directly evoked specific aspects of the implementation process, such as the nurses’ expectations for televisits (asked as follows: “What expectations did you have regarding telemedicine and televisits [note: before the implementation process]?”), the questions themselves predefined the main themes. The coding was done manually within different columns in Microsoft Excel (Microsoft Corp). The responses of the interview partners were divided into sections with different ideas, each collected in separate rows. In a second step, inductive coding was done by defining labels created based on the data. These were then regrouped into categories based on recurring patterns. The labels and the themes were translated into English. To present the data, the different categories under each main theme were condensed into key points or brief sentences.

Follow-Up Questionnaire

In June 2023, a follow-up questionnaire (Multimedia Appendix 5) was distributed via a web-based survey tool (UmfrageOnline; enuvo GmbH) to all the nursing staff in NH 1, in which long-term implementation of televisits had been achieved. For comparing pre- and postimplementation results, nurses were tasked with evaluating the time spent on televisits and the impact of implementing televisits on residents’ medical care, mirroring the approach used in the initial questionnaire. In multiple-choice questions, the nurses also reported the impact on nursing practice. Several other aspects, such as the interaction with the GP, the usability of the telemedical system, and the potential of televisits with physicians from medical specialties other than family medicine, were assessed. Only the main results are presented in this paper.

Ethical Considerations

This study was approved by the ethics committee at the Faculty of Medicine of the Rheinisch-Westfälische Technische Hochschule Aachen (EK 23-178).

Results

Questionnaire on Baseline Expectations and Competencies

Before the implementation of televisits, nurses rated both their computer competencies and their clinical skills in taking vital parameters and examining a resident using medical diagnostic devices as good, with slightly better levels in NH 2 (computer competencies: median 2.0, IQR 2.0-3.0, and n=11 in NH 1 vs median 1.5, IQR 1.0-2.0, and n=8 in NH 2, P=.06; diagnostic devices: median 2.0, IQR 2.0-2.0, and n=11 in NH 1 vs median 1.5, IQR 1.0-2.0, and n=8 in NH 2, P=.10).

With regard to baseline expectations, slightly more than half (7/11, 64% of respondents in NH 1 and 4/8, 50% of respondents in NH 2; P=.66) of the nurses expected televisits to improve the residents’ care, 4 nurses expected the positive and negative effects of televisits to be equal (0/11, 0% in NH 1 and 4/8, 50% in NH 2; P=.02), and 20% of the total nurses (4/11 in NH1 and 0/8, 0% in NH2, P=.1) agreed with neither of these 2 statements (Figure 3). In total, 4 nurses (3/11, 27% in NH 1 and 1/8, 13% in NH 2; P=.60) declared that televisits should only be performed in exceptional cases (see the raw data in Multimedia Appendix 2), 1 (25%) of whom expected that delivering televisits to older residents would be difficult and 2 (50%) of whom did not expect to see positive effects from televisits on the residents’ care (Figure 3). The effect of implemented televisits on the nursing workload was estimated to be neutral by the vast majority in both NHs (9/11, 82% in NH 1 and 6/8, 75% in NH 2; P>.99) and to be time saving or time consuming by 3 (1/11, 9% in NH 1 and 2/8, 25% in NH 2) and 1 (1/11, 9% in NH 1 and 0/8, 0% in NH 2) respondents, respectively, with no significant differences between the NHs (P=.55 and P>.99, respectively).
The initial training sessions on the diagnostic devices were rated as very good to good, and the software instructions were rated as good to satisfactory.

**Televisit Documentation Protocols**

A total of 30 televisits in NH 1 and 19 televisits in NH 2 performed between May 2020 and January 2021 were documented by 10 and 6 nurses, respectively, averaging 3.00 and 3.16 televisits per nurse, respectively. Of the total of 49 televisits, 36 (73%; 23/30, 77% in NH 1 and 13/19, 68% in NH 2; \( P = .53 \)) were used for planned routine assessment. In 13 cases (7/30, 23% in NH 1 and 6/19, 32% in NH 2; \( P = .53 \)), the televisits were scheduled at short notice for urgent assessments of residents who had acute medical presentations and issues. In NH 2, all the medical devices provided by the telemedical system were systematically used in every single televisit. In NH 1, the televisits included the use of 1 medical device in 50% (15/30) of cases. In 30% (9/30) of cases, the system was used for videoconferencing without further use of diagnostic devices. In total, 2 or 3 medical devices were only used in 13% (4/30) and 7% (2/30) of televisits, respectively. The blood pressure meter was the PoC device used most often for televisits (16 times), followed by the electronic stethoscope (8 times) and the 1-canal ECG (3 times; Table 1). During these first televisits, the nurses rated their competencies in handling the software and the medical devices as good (Table 2). In an open-ended question asking for possible improvements, they mentioned the need for further training in half (8/16, 50%) of the answers, followed by technological (6/16, 37%) and organizational (2/16, 13%; Multimedia Appendix 6) issues.
Table 1. Use of point-of-care (PoC) medical devices in nursing home (NH) 1 within the initial testing and familiarization period. Data presented for a subset of n=21 televisits where one or multiple PoC devices were used. In 9 additional televisits documented in NH 1, no PoC devices were used. The data for NH 2, where measurements were systematically taken using all the PoC devices in every televisit, are not presented (n=19).

<table>
<thead>
<tr>
<th>Medical device</th>
<th>Uses, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BP&lt;sup&gt;a&lt;/sup&gt; meter</td>
<td>16 (53)</td>
</tr>
<tr>
<td>Blood glucose meter</td>
<td>2 (7)</td>
</tr>
<tr>
<td>1-canal ECG&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3 (10)</td>
</tr>
<tr>
<td>Electronic stethoscope</td>
<td>8 (27)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of devices</th>
<th>Uses, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No diagnostic devices</td>
<td>9 (30)</td>
</tr>
<tr>
<td>1 device</td>
<td>15 (50)</td>
</tr>
<tr>
<td>2 devices</td>
<td>4 (13)</td>
</tr>
<tr>
<td>3 devices</td>
<td>2 (7)</td>
</tr>
</tbody>
</table>

<sup>a</sup>BP: blood pressure.

<sup>b</sup>ECG: electrocardiogram.
Table 2. Self-reported competencies in handling the telemedical software and the point-of-care medical devices as assessed by the nurses themselves. *P* values are given for the Mann-Whitney U test comparing the results in nursing homes (NHs) 1 and 2. Scale from 1.0 to 6.0, with 1.0 corresponding to very good and 6.0 to deficient.

<table>
<thead>
<tr>
<th>Point-of-care device and population</th>
<th>Number of ratings, n</th>
<th>Values, median (IQR)</th>
<th><em>P</em> value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Camera</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH 1</td>
<td>20</td>
<td>1.0 (1.0-2.0)</td>
<td>__a</td>
</tr>
<tr>
<td>NH 2</td>
<td>19</td>
<td>2.0 (2.0-2.0)</td>
<td>.02</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>2.0 (1.0-2.0)</td>
<td>__</td>
</tr>
<tr>
<td><strong>Blood pressure meter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH 1</td>
<td>14</td>
<td>1.0 (1.0-2.0)</td>
<td>__</td>
</tr>
<tr>
<td>NH 2</td>
<td>19</td>
<td>2.0 (1.0-2.0)</td>
<td>.73</td>
</tr>
<tr>
<td>Total</td>
<td>33</td>
<td>2.0 (1.0-2.0)</td>
<td>__</td>
</tr>
<tr>
<td><strong>Blood glucose meter</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH 1</td>
<td>1</td>
<td>1.0 (1.0-1.0)</td>
<td>__</td>
</tr>
<tr>
<td>NH 2</td>
<td>19</td>
<td>2.0 (1.0-2.0)</td>
<td>.31</td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>2.0 (1.0-2.0)</td>
<td>__</td>
</tr>
<tr>
<td><strong>1-canal ECG</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH 1</td>
<td>3</td>
<td>2.0 (1.5-2.5)</td>
<td>__</td>
</tr>
<tr>
<td>NH 2</td>
<td>19</td>
<td>2.0 (1.0-2.0)</td>
<td>.47</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>2.0 (1.0-2.0)</td>
<td>__</td>
</tr>
<tr>
<td><strong>Electronic stethoscope</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH 1</td>
<td>8</td>
<td>2.0 (1.75-3.0)</td>
<td>__</td>
</tr>
<tr>
<td>NH 2</td>
<td>19</td>
<td>2.0 (1.0-2.0)</td>
<td>.09</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>2.0 (1.0-2.0)</td>
<td>__</td>
</tr>
<tr>
<td><strong>Software</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH 1</td>
<td>29</td>
<td>2.0 (1.0-3.0)</td>
<td>__</td>
</tr>
<tr>
<td>NH 2</td>
<td>19</td>
<td>2.0 (1.0-2.0)</td>
<td>.13</td>
</tr>
<tr>
<td>Total</td>
<td>48</td>
<td>2.0 (1.0-2.0)</td>
<td>__</td>
</tr>
</tbody>
</table>

*a*Not applicable.

*b*ECG: electrocardiogram.

**Systematic Analysis of the First Implementation Phase**

**Initial Organization of Televisits**

In the first phase of implementation, televisits were organized in the same way as on-site home visits, where a physician sees several residents in an NH while moving between the different residents’ rooms. The remote physician working in the GP office connected via the telemedical system to a nurse in the NH standing next to a resident to be consulted. The televisits had been synchronous at the beginning—thus, the GP in the office as well as the nurse and the resident in the NH both connected to the telemedical interface at the same time and no previous assessment of the resident such as taking of the vital parameters had been conducted beforehand. When the GP requested measurements, the nurses used the medical devices for taking vital signs during the ongoing televisits. Meanwhile, the physician connected to the telemedical system in the GP office via internet waited until completion before deciding which steps to take next. When one televisit had been completed, the nurse on-site moved the telemedical device (with the physician connected via internet) from one resident’s room to another. During these transfer times, the GP also waited. We called this modality *synchronous televisits in the modality of web-based home visits* (Figure 4). During the first implementation period, the issues evoked by the NHs and addressed by the steering committee essentially concerned signal stability, specific software features, and the use of the medical devices. Issues on the side of the GP such as the billing of telemedical visits were also addressed. Televisits were performed regularly in both NHs until the easing of lockdown restrictions in Germany in June 2021.
Figure 4. Synchronous televisits in the modality of web-based home visits by a remote general practitioner (GP) in a GP office (1) connected via a telemedical system (2) with a nurse in the nursing home next to a resident (3).

Descriptive Presentation of Encountered Challenges and Adopted Control Measures

Health Care Context at the Time of the Systematic Analysis
When the German health care system and society were in transition between the pandemic and the postpandemic situation, the steering committee systematically analyzed in August 2021 the experiences of the first implementation phase. At that point, the last German COVID-19 restrictions were over, and normal home visits were possible again. Resistance to change was rising in the HCPs, who reported a lot of challenges and issues. Three main barriers to long-term adoption were identified: (1) conflicting workflows between the NHs and the GP offices, (2) the lack of time efficiency of the televisits, and (3) perceived uncertainty in using the technology among the nursing staff.

Conflicting Workflows
The daily routines and workflows were different in the NHs and in the GP offices. Finding suitable moments for televisits was difficult. For example, noon was a time when there were no consultations in the GP office, and thus, it was ideal for televisits from the GP’s perspective. However, in the NHs, all the staff were busy serving and delivering lunch. Initially, televisits were scheduled at varying times every week, resulting not in the desired flexibility but in organizational stress and discomfort. This was addressed by agreeing on fixed weekly time slots for the routine televisits. With regard to unplanned televisits in the case of acute medical presentations, the workflows were also initially conflicting. The NHs did not bundle calls to the GP office. As a consequence, the workflow of the GP office was interrupted multiple times a day when 2 or more visits had to be planned. During the consultation hours in the GP office, phone calls were particularly disruptive because the GP was seeing other patients at that moment. To avoid disruption due to phone calls, the GP office and the NH grouped requests. Phone calls were completely abandoned except for extreme emergencies. The NH sent a fax containing a list with the residents proposed for consultations along with the reasons for consultation and some contextual information. On the basis of this information, the GP then responded with a fax indicating the vital parameters that should be assessed before the actual consultation (the organization of televisits was changed for an asynchronous modality with vital parameters being taken beforehand; see the following section).

Lack of Time Efficiency of the Televisits
The televisits, highly valued during the pandemic phase, were perceived as overly time consuming compared to normal home visits when the latter were regularly possible again. As mentioned previously, the workflows were interrupted during the transit times and when vital signs were taken. This was perceived by the GPs as a significant loss of time they could not spend with other patients in their offices. For the GPs, the time spent on televisits needed to be reduced so that the televisits were a real benefit to them compared to on-site home visits. As these challenges were mainly linked to the synchronous modality of televisits, the organization was completely changed for the modality of asynchronous televisits using a web-based waiting room (Figure 5). In this approach, the vital signs of the residents...
to be seen are taken beforehand by the nurses on-site. Then, with all diagnostic measures completed and documented, the televisits are performed at a scheduled time. The residents to be consulted physically wait in a waiting room in the NH. They also appear in a web-based waiting room in the GP’s telemedical interface. The televisits take place in a dedicated room where the telemedical cart is used while stationary. Instead of moving the telemedical cart, the residents enter the televisit room one at a time as they would a consultation room in a GP office (Figure 5). This greatly improved the time efficiency of the televisits. To avoid losing time spent with doubled-up documentation, interfaces among the telemedical system, the GP, and the NH documentation systems were created. This made it possible to directly export the medical documentation from the TeleDoc software to the other programs.

**Figure 5.** Asynchronous televisits using a web-based waiting room. Step 1: taking of the vital parameters of the residents to be seen. Step 2: asynchronous televisits at a scheduled time using a web-based waiting room for the general practitioner (GP; 1), a physical waiting room in the nursing home (NH; 2), and a defined televisit room (3). The control actions were the creation of the web-based waiting room (A), training of the nursing staff (B), allowing for asynchronous documentation in the software (C), and interfaces to the existing programs (D and E).

**Step 1: preparation of the televisits**

1. Resident room 1
2. Resident room 2
3. Resident room 3

**Step 2: asynchronous televisits using a web-based waiting room**

1. GP office with a web-based waiting room
2. Waiting room in NH
3. Televisit room

**Perceived Uncertainty in Using the Technology Among the Nursing Staff**

Moreover, the nursing staff also requested more intense training. To this end, NH 1 and the cooperating GP office jointly hosted a medical student for 4 weeks, during which the student supported the training of the nursing staff. He trained nurses in the use of the PoC devices, helped prepare televisits, and accompanied them during these visits. During these 4 weeks, daily televisits were performed. Finally, long-term implementation was reached in NH 1, with regular, most of the time weekly, televisits between the GP and
NH 1 from August 2021 until the end of follow-up in March 2023. A total of 163 televisits were performed during this period of 20 months. The televisits did not replace the preceding rhythm of on-site home visits but were performed in addition. In NH 2, further efforts to implement televisits after the COVID-19 period remained unsuccessful throughout the follow-up period.

**Interview Findings**

**Selection and Characteristics of Interview Participants**
In both NHs, it was planned that the first 5 nurses who volunteered would be included. In NH 2, many nurses considered their experience with telemedicine to be too little and too far away, so no one else wanted to participate after the first 3 nurses were recruited. As NH 2 did not achieve long-term implementation and had not regularly performed televisits since August 2021, additional interviews were unlikely to reveal new themes, insights, or perspectives. Thus, data saturation had already been reached after these 3 interviews. The interviewed nurses were predominantly male (5/8, 62%), with a median age of 26 years. The youngest and oldest participants were aged 26 and 63 years, respectively. The nurse managers were all male (3/3, 100%), with the youngest and oldest participants being aged 30 and 56 years, respectively. The interviews took, on average, 33 minutes for the nurses and 40 minutes for the nurse managers.

**Interview Findings Concerning the Implementation Phases**
In the follow-up interviews, the nurses and nurse managers stated that their main expectation before the implementation of televisits was to enhance residents’ access to physicians, particularly beyond the regular GP consultation hours. The nurse managers further expected a decrease in psychological distress in their staff, along with enhanced legal protection through telemedical expertise during emergency situations with acute health-related deteriorations in their residents.

As regards the subsequent implementation, the interview findings are consistent with the analysis of the steering committee. Nurses in both NHs agreed that televisits were seen as useful during the COVID-19 pandemic and had offered manifest benefits for the residents. Televisits addressed the specific need for thorough medical assessment in the context of the then current contact restrictions. However, the benefits of the televisits seemed less identifiable for most of the nurses of both NHs after the end of the contact restrictions given that normal home visits were possible again and because the telemedical system did not provide the residents with better access to the health care system in terms of other medical specialties or access to physicians outside of the normal office hours. In total, 25% (2/8) of the nurses explicitly saw a benefit of televisits with the GP for residents with acute medical presentations outside the hours of the on-site home visits.

Nurses in NH 1, where long-term implementation was achieved, all agreed that “long-term benefits of televisits became apparent only after many repeated adjustments and training sessions.” The nurses stated that the system became more user-friendly and easier to handle. Repeated use led to more routine use and more self-confidence in the televisits. For this, the repeated training and accompaniment in the televisits were very important for the nurses.

When asked about the factors that had contributed to the success of long-term implementation, the nurses underlined the adoption of asynchronous visits and the transformation of the workflows to a modality with a web-based waiting room. They also mentioned the crucial role of the nurse manager’s and the GP’s personal commitment in ensuring the long-term implementation of televisits. The fact that there was a clear benefit for the residents also added to the motivation.

**Interview Findings Regarding the Change in Nursing Practice Due to Televisits**
The 3 interviewed NH and nurse managers and 88% (7/8) of the interviewed nurses retrospectively evaluated the overall experience as positive. However, one nurse felt that the televisits were too impersonal and described difficulties in handling the telematic system and in performing televisits for people with dementia. This nurse preferred in-person home visits.

The nurse managers stated that televisits reduced the psychological distress experienced by their nursing staff, improved the clinical skills of their employees, and led to greater employee satisfaction during the COVID-19 pandemic (in both NHs) and beyond (for NH 1 in achieving long-term implementation). The interviewed nurses mostly acknowledged the same effect on their clinical skills, but most denied psychological distress and insecurities in their work environment. While most of the nurses stated that asynchronous televisits compared to home visits translated to substantial time savings for the GP, the opinions varied regarding the effect on the nursing workload. Televisits were rather perceived as additional workload when performed for routine visits. In the case of unscheduled visits with GPs, some nurses praised the televisits as time saving, whereas other nurses considered telephone calls to be the faster solution. With regard to consultations outside GP hours, the nurses consistently declared that televisits would allow for a faster medical assessment and reduce nursing workload compared to calling the in-person out-of-hours GP service, for which a considerable waiting time is usual. On top of the organizational and technical challenges regarding the cooperation between the GP offices and the NHs, the nurses described personal challenges in adapting to change, using the technology, and developing effective skills for video communication. Some nurses saw positive effects of televisits on the nurses’ self-image and with regard to legal certainty. Training in simulated televisits and support from nursing peers in a tandem approach were considered adequate to offer continuous training after the initial familiarization phase (Textbox 1).

https://aging.jmir.org/2024/1/e55471
Textbox 1. Summarized interview findings regarding relevant topics raised during the implementation process.

**Psychological distress in situations with acute deterioration of a resident’s health status (need for prompt medical assessment)**

- A total of 62% (5/8) of the nurses denied experiencing psychological distress in these situations.
- A total of 38% (3/8) of the nurses described psychological distress in these situations.
- Access to televisits can provide nurses with emotional security and reduce psychological stress.

**Clinical skills of the nurses**

- Most nurses saw improvements in their clinical skills, mainly due to more diagnostic equipment being available in the nursing home (NH), increased awareness of recognizing changes in residents’ conditions, and a more active participation in the residents’ medical care.
- Nurses reported higher confidence in the interactions with physicians since the implementation of televisits.

**Nursing workload**

- Asynchronous televisits saved time for the general practitioner (GP).
- Routine televisits seemed to be associated with a slight increase in workload.
- For unscheduled visits, some nurses considered the televisits to be time saving. However, for other nurses, telephone calls seemed to be faster.
- Televisits can be time saving for the nursing staff provided that a telemedical system is always available on standby (no loss of time due to booting), the nurses are experienced in televisits, and a physician is quickly available (direct medical assessment).
- A telemedical service with 24/7 access to a GP was seen as an advantage over the current out-of-hours GP service and would save time for the NHs (reduced waiting time—obsolete telephone waiting line and transfer time of the physician on duty).

**Challenges in performing televisits at the beginning**

- Technical challenges (software bugs, usability, and user experience aspects) and change management (for generating willingness in the nursing staff and the physicians) were seen as challenging at the beginning.
- Collaboration with the physicians, especially in scheduling televisits, and ensuring adequate staffing and support for the televisits were identified as the biggest organizational challenges.

**Communication in a televisit between the nurse, the resident, and the physician**

- It was completely unproblematic for one nurse and difficult and uncomfortable in all televisits for another.
- Most nurses reported challenges in this new setting at the beginning.
- Acoustic problems made televisits difficult for residents with hearing impairments, particularly those who partly relied on lipreading.
- By ensuring a calm and friendly environment and facilitating communication, televisits were possible for residents with dementia.
- Communication is likely to be easier with the next generations of residents in NHs who are already more familiar with modern communication technology.

**Professional identity and self-image of the nurses**

- Nurses understood televisits as a logical and upcoming innovation arriving in the nursing profession in the context of general progress in technology and digitalization.
- However, they did not think that televisits elevate the nursing profession.
- One nurse felt that his self-image in the interaction with the physician improved by experiencing close cooperation and teamwork in televisits.

**Medico-legal aspects of the telemedical documentation**

- Main advantage: written prescriptions and medication schedules established by the physician and serving as legal documents are available faster in the NHs.
- However, nurses did not consider the legal certainty that a televisit provides to be higher than that of telephone calls and their subsequent documentation.

**Cooperation between GPs and NHs**

- In both NHs, the cooperation with the physicians was perceived as already good or very good before the introduction of televisits.
- The implementation of televisits did not change the cooperation from the perspective of the nurses.
Follow-Up Questionnaire: Time Expenditure and Change in Nursing Practice Due to the Televisits

The follow-up questionnaire showed that most of the responding nurses performed televisits ≤5 times throughout the study period and only some nurses did so >10 times (>10 times: 4/19, 21%; 5-10 times: 4/19, 21%; 1-5 times: 11/19, 58%; median 1-5 times). Half (9/18, 50%) of the nurses agreed with the statement that televisits were associated with an additional burden related to uncertainty in using the new technology. The nurses rated the time expenditure for performing televisits significantly higher at the end of the follow-up than they did before starting the first televisits. At both timepoints, only very few nurses considered the televisits to be time saving for them (1/11, 9% before implementation, vs 3/18, 17% after implementation; \( P < .001 \)). While the nurses did not expect the televisits to be time consuming in the initial assessment (1/11, 9%), the assessment was significantly different after the implementation process, where the vast majority of the nurses rated them as time consuming (12/18, 67%; \( P = .006 \)). The relative percentage of nurses expecting a neutral effect on the nursing workload was accordingly lower after implementation (3/18, 16%) than before (9/11, 82%; \( P = .001 \)). Compared to the assessment done before the intervention, the evaluation of the impact of televisits on the care of the residents remained unchanged, with half (10/19, 53%) of the nurses estimating a positive effect, one-third (6/19, 31%) estimating a neutral effect, only 1 person estimating a negative effect (1/19, 5%), and two persons with missing answers (2/19, 11%; \( P > .99 \) for the positive effect and \( P = .71 \) for the neutral effect when compared to the preimplementation assessment; see the Evaluation of the Initial Testing and Familiarization Period Up Until January 2021 section). Most nurses fully or generally agreed that televisits were entirely suitable for assessing the residents and for initiating treatment decisions in the case of GP-related medical queries (full approval: 5/19, 26%; predominant approval: 11/19, 58%; mild disagreement: 2/19, 11%; full disagreement: n=0; 1/19, 5% missing answers). As in the interviews, the quantitative evaluation also showed further positive effects of the televisits in addition to improving the clinical and technological competencies of the nurses. These include a sense of empowered participation in the implementation process for about every fourth responding nurse (5/18, 28%), higher professional recognition from the GP and the residents for approximately every third responding nurse (each 7/18, 39%), and a decrease in psychological distress and greater legal certainty for more than half of the respondents (10/18, 56%; Multimedia Appendix 7).

Discussion

**Principal Findings**

**Key Drivers for Successful Implementation**

This study evaluated the implementation of televisits during the COVID-19 crisis in 2 NHs with a follow-up in the postpandemic setting. While both NHs achieved short-term implementation within the pandemic phase, only 1 NH attained long-term implementation. This was achieved in an action-based approach after control measures had been implemented following an analysis of the barriers to further implementation. Adapting the telemedical system and changing the workflows especially to asynchronous televisits with a web-based waiting room were key drivers for successful implementation.

There was no difference in the initial evaluation between the 2 NHs explaining why long-term implementation was only achieved in NH 1. It is noteworthy that, during the first implementation phase, NH 1 adopted a more pragmatic resident-centered approach using only the PoC devices required for the individual patients as shown in the documented televisits. In contrast, NH 2 systematically used all the devices for training purposes in their initial televisits. Moreover, there is a common perception that older nursing staff may possess lower levels of digital literacy, which in turn might render them more resistant to the adoption of new technologies. Interestingly, the staff in NH 1 was older than that in NH 2. Several factors may explain why, contrary to this perception, long-term implementation was achieved in NH 1 with the older nurses.

First, NH 2 simultaneously engaged in a second change process after the end of the COVID-19 pandemic when it changed its NH documentation software in November 2021. At this point, the nursing staff of NH 2 were extremely challenged by adapting to the new software. This competing project was more prioritized than continuing to implement televisits. Second, the interviews revealed that the nurses felt that the personal engagement of the nurse manager and the GP was very important for achieving long-term implementation. This barrier is consistent with findings in the literature indicating that gaining leadership and clarifying roles are important for driving implementation processes [32,34]. Competencies and criteria for leaders in implementing digital health care should be clearly defined in the future to enhance the implementation of telemedicine and digital health [41,42]. Third, NH 1 developed a specific training strategy with the medical student who helped train the nursing staff for 4 weeks. This also contributed greatly to the nurses’ compliance and subsequent implementation.

**Importance of Considering the User Perspective**

The initial vision of how the televisits would be delivered differed significantly from how they were finally integrated by the HCPs to achieve long-term adoption. At the beginning of
the transformation process, everybody involved (developers, GPs, nurse managers, and nurses) envisioned a modality that mirrored the workflow of a normal GP home visit (synchronous televisits in the modality of web-based home visits). However, another modality, that of “asynchronous televisits using a web-based waiting room,” proved to be more suitable and efficient for routine consultations. The conceptual shift from web-based home visits to the web-based waiting room led to profound organizational modifications and practice changes. Both the participating nurse managers and the developers of the telemedical system stated that they could not have anticipated the web-based waiting room modality before starting the implementation. This highlights the importance of considering the users’ perspective in all phases of development of eHealth solutions, including once a new system is made market available. This approach of broadly involving “a wide range of stakeholders in the entire development process, including especially end-users—patients and physicians” is referred to as “co-creation” [43]. In this study, the HCPs acted as cocreators for developing the asynchronous televisits using a web-based waiting room. They reported their in-use experience with the telemedical system to the steering committee, which triggered discussion on change within the committee. The telemedical system and the workflows were then adapted to allow for asynchronous televisits. This resulted in an improved telemedical system and, thus, a more valuable product. As it provides greater usability for the HCPs, their acceptance of the system and the implementation process increases. The CM approach in this study allowed for this interaction between end users and developers and additionally provided an organized platform for identifying ongoing challenges and steering the transformation process in the right direction. Imposing overly rigid application frameworks for new digital innovation in health care may be a major contributor to the failure of health innovations to be transferred into routine care.

**Change in Nursing Practice Due to Televisits**

In NH 1, the nurses were positive about the long-term implementation and acknowledged improvements in their clinical skills and technological competencies. In addition, the televisits improved their self-image and their recognition as nurses. These positive results were found in the interviews and confirmed in the follow-up questionnaire. When asked directly in the interviews, only a minority of the nurses reported experiencing physical distress when residents show acute medical presentations requiring prompt medical assessment. However, most agreed in the quantitative analysis that access to televisits was perceived as relieving and reducing emotional distress. The lack of consistency here is probably due to response bias, with nurses responding more honestly in an anonymous questionnaire than in a face-to-face interview.

Televisits were perceived as time saving for the GP but were associated with an additional technological burden and a slight increase in time spent by the nursing staff. The latter was not expected by the nurses when they were assessed before the implementation process. Interestingly, this observation is in line with those of other studies [44-46], which also show that televisits in NHs are associated with an increased workload for the nursing staff. Although most nurses in this study expected a neutral effect on workload rather than an increase, the staff in NH 1 adapted to televisits even though they increased their workload. This is a good indication that the other effects are seen as positive and valuable.

The taking of vital parameters using medical diagnostic devices or recording on an ECG are not among the normal tasks of a geriatric nurse in Germany. This means that there is unlikely to be an increase in the overall workload but rather a shift from GP activity to tasks performed by geriatric nurses. The GPs can use the time they save for other patients. This can be beneficial in the context of a shortage of GPs, especially in rural areas. Given the overall positive effects, health insurers and public health managers should consider providing financial support for the implementation of televisits in NHs. In particular, the initial setup and training should be financially supported to maintain high motivation for further implementation.

This work logically builds on and complements the group’s previous research, which demonstrated the technical feasibility and usefulness of integrating PoC diagnostic devices into video consultations for the assessment of older adult patients in NHs [38]. This study shows that there is no need to take all the vital parameters of the residents in every consultation. In more than three-quarters of the televisits (24/30, 80%), no or one single measurement was sufficient. Access to the right diagnostic device at the right time for the right resident with the right physician connected via telemedicine is essential for personalized care and to avoid unnecessary HAs. In line with the idea of effective workforce management, it does not make sense to ask nurses to take all the vital parameters but rather to decide in advance (ie, when planning the televisit) which vital parameters should be taken.

**Limitations and Future Research**

This study has certain limitations for the interpretation and generalizability of the results. The implementation of televisits in NHs was only evaluated in 2 NHs, and 1 failed to achieve long-term implementation. Therefore, generalization to other NHs may depend strongly on their individual organization and health care environment. However, this study shows that an action-based CM approach piloted by an interdisciplinary steering committee can allow for successful implementation. Future studies in the field of IS may add valuable insights to determine more objective criteria for the success of telemedical implementation efforts. Another major limitation is the relatively small number of televisits in the documentation protocols. However, this was sufficient to show that it is not useful to systematically take measurements using all diagnostic devices integrated into the system. This study was not designed to assess how the residents felt about being involved in promoting telecare. However, this is an important research question that should be evaluated in other studies. Furthermore, this study investigated the implementation of one technology—televisits—in NHs. Other digital technologies such as NH managing software, wearables, and home automation systems may also improve workforce management, enhance the quality of care, and provide a better living environment in NHs. As some authors believe that all these technologies are likely to be integrated into so-called “smart
NHs,” possibly replacing conventional NHs in the decades to come [47-49], the implementation processes of these technologies should be specifically investigated as generalizability from this study is very limited.

Future research should focus on training and staffing concepts for nursing. First, performing televisits is a new approach to care and needs specific education. While the need for specific curricular training in telemedicine has already been identified for medical students [50-53], it should also be integrated into nursing education. Second, a notable finding in the follow-up questionnaire was that most televisits were conducted by a small number of nurses with some performing a significant number whereas most only conducted a few (>10 televisits: 4/19, 21%; <10 televisits: 15/19, 79%). Therefore, future research should explore whether certain nurses should specialize in televisits within a nursing facility whereas less qualified nurses or nursing assistants could focus on other nursing tasks. Maybe this would lead to workforce optimization. In theory, telemedically advanced nurses could also use telemedicine for telenursing, that is, for advising and supporting nursing assistants via telemedical solutions.

Conclusions
An action-oriented CM approach made it possible to achieve long-term implementation of televisits in NH 1. Abandoning synchronous televisits in favor of asynchronous ones improved the workflows and was a critical facilitator of long-term implementation. Strong leadership, as well as sustained training of the nurses, also contributed to this success. The implementation of televisits had positive effects on the HCPs. They experienced an improvement in their clinical skills and a higher professional recognition. According to the nurse managers, their psychological distress also decreased. Performing televisits did not save time for the nursing staff in comparison to scheduling and assisting a GP in in-person home visits. Instead, the nurses believed that televisits led to a slight increase in their time spent on organizing the assessment of the patient. CM aspects must be considered to achieve long-term implementation.

Acknowledgments
This study was conducted within the AIDA project (Arbeitsentwicklung in der Altenpflege) funded by the European Regional Development Fund 2014 to 2020 of the European Union and by the state of North Rhine-Westphalia, Germany. The authors acknowledge and thank the nursing homes “Stiftung Evangelisches Alten- und Pflegeheim Gemünd” (Schleiden, Germany) and “St. Gereon Seniorendienste gGmbH” (Hückelhoven, Germany), as well as Ekaterini Karidas-Moitzheim, Ines Ferber, Arno Brauckmann, Gerd Palm, and Jan Fabry for their cooperation and support.

Authors' Contributions
TM and AF conceptualized and designed the study. SV organized the initial training of the health care providers. HG and PFF provided advice on the qualitative research. AF, MC, and PFF obtained research funding. TM collected, managed, and analyzed the data. TM interpreted the findings and drafted the manuscript. All authors contributed substantially to its revision. All authors approved the published version of the manuscript. TM takes responsibility for the paper as a whole.

Conflicts of Interest
MC is the chief executive officer and cofounder of Docs in Clouds TeleCare GmbH, a company that purchases telemedical solutions including the TeleDoc system, which was used in this study. AF worked part time as chief medical officer for Docs in Clouds TeleCare GmbH until March 2023. All other authors declare that they have no other conflicts of interest.

Multimedia Appendix 1
Project diary.
[PDF File (Adobe PDF File), 1430 KB - aging_v7i1e55471_app1.pdf ]

Multimedia Appendix 2
Excel (Microsoft Corp) sheet data from project diaries.
[XLSX File (Microsoft Excel File), 25 KB - aging_v7i1e55471_app2.xlsx ]

Multimedia Appendix 3
COREQ (Consolidated Criteria for Reporting Qualitative Research) reporting guidelines.
[PDF File (Adobe PDF File), 186 KB - aging_v7i1e55471_app3.pdf ]

Multimedia Appendix 4
Interview guide.
[PDF File (Adobe PDF File), 295 KB - aging_v7i1e55471_app4.pdf ]
Multimedia Appendix 5
Follow-up questionnaire.

Multimedia Appendix 6
Comments of the nurses regarding improvement potential for further televisits. The number of times a specific topic was mentioned is indicated in brackets.

Multimedia Appendix 7
Nurses’ self-assessment of the positive effects of implementing televisits in nursing home 1. Agreement with the individual aspects shown for 18 nurses; n=1 respondent in the postevaluation period did not select an answer.

References


Abbreviations

CM: change management
COREQ: Consolidated Criteria for Reporting Qualitative Research
ECG: electrocardiogram
GP: general practitioner
HA: hospital admission
HCP: health care professional
IS: implementation science
NH: nursing home
PoC: point-of-care
Original Paper

Tailored Self-Management App to Support Older Adults With Cancer and Multimorbidity: Development and Usability Testing

Sang-Wha Sien1*, BCScpE, MSHCDE; Francis Kyerepagr Kobekyaa1*, BA, MN; Martine Puts2, BSCN, BScPT, PhD; Leanne Currie3, BN, MScN, PhD; Margaret Tompson4, PhD; Penelope Hedges5, BA; Joanna McGrenere1, BSc, MSc, PhD; Caroline Mariano5, BSc, MD; Kristen R Haase3, BN, MA, PhD

1Department of Computer Science, University of British Columbia, Vancouver, BC, Canada
2Lawrence S Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada
3School of Nursing, University of British Columbia, Vancouver, BC, Canada
4Saskatchewan Center for Patient Oriented Research, Saskatoon, SK, Canada
5BC Cancer, Vancouver, BC, Canada

* these authors contributed equally

Corresponding Author:
Kristen R Haase, BN, MA, PhD
School of Nursing
University of British Columbia
T201-2211 Wesbrook Mall
Vancouver, BC, V6T 2B5
Canada
Phone: 1 604 822 0979
Email: kristenhaase@gmail.com

Abstract

Background: Globally, cancer predominates in adults aged older than 60 years, and 70% of older adults have ≥1 chronic condition. Cancer self-management interventions can improve symptom management and confidence, but few interventions target the complex needs of older adults with cancer and multimorbidity. Despite growing evidence of digital health tools in cancer care, there is a paucity of theoretically grounded digital self-management supports for older adults. Many apps for older adults have not been co-designed with older adults to ensure that they are tailored to their specific needs, which would increase usability and uptake.

Objective: We aim to report on the user evaluations of a self- and symptom-management app to support older adults living with cancer and multimorbidity.

Methods: This study used Grey’s self-management framework, a design thinking approach, and involved older adults with lived experiences of cancer to design a medium-fidelity app prototype. Older adults with cancer or caregivers were recruited through community organizations or support groups to participate in co-designing or evaluations of the app. Data from interviews were iteratively integrated into the design process and analyzed using descriptive statistics and thematic analyses.

Results: In total, 15 older adults and 3 caregivers (n=18) participated in this study: 10 participated (8 older adults and 2 caregivers) in the design of the low-fidelity prototype, and 10 evaluated (9 older adults and 1 caregiver) the medium-fidelity prototype (2 older adults participated in both phases). Participants emphasized the importance of tracking functions to make sense of information across physical symptoms and psychosocial aspects; a clear display; and the organization of notes and reminders to communicate with care providers. Participants also emphasized the importance of medication initiation or cessation reminders to mitigate concerns related to polypharmacy.

Conclusions: This app has the potential to support the complex health care needs of older adults with cancer, creating a “home base” for symptom management and support. The findings from this study will position the researchers to conduct feasibility testing and real-world implementation.

(JMIR Aging 2024;7:e53163) doi:10.2196/53163

https://aging.jmir.org/2024/1/e53163
KEYWORDS
cancer; aging; self-management; usability testing; design thinking; design; oncology; develop; development; usability; gerontology; geriatric; geriatrics; older adult; older adults; elder; elderly; older person; older people; ageing; mHealth; mobile health; app; apps; application; applications; symptom; symptoms; comorbidity; comorbidities; comorbid; multimorbidity; multimorbidities; co-design

Introduction
Cancer morbidity and mortality increase with age for most cancers [1,2]. With the rapidly aging population, the number of older adults (≥65 years) with cancer is estimated to double globally by 2035 [1] and triple in those aged ≥80 years in the next decades [3]. In Canada, 2 in every 5 older adults aged ≥70 years are diagnosed with cancer, accounting for 28.2% of all deaths [4]. Approximately 70% of older adults with cancer have pre-existing illnesses that occur with aging [5]. Having multiple conditions alongside cancer can lower one’s functional and cognitive status, increase the likelihood of treatment complications, and negatively impact health outcomes [1,3,5].

Given the possible deleterious side effects of cancer-related treatments, older adults and their families require self-management support during the cancer care trajectory [6,7]. Self-management refers to the ability to manage treatment effects and psychosocial changes arising as a result of illness [7]. Within the context of cancer, self-management refers to one’s ability to manage the effects of diagnosis and treatment [7]. Self-management encompasses the capacity to manage the symptoms, treatment, physical and psychosocial consequences, and lifestyle changes inherent in living with cancer [6]. Self-management support programs are often tailored to the needs and abilities of patients and their relatives [6] and comprise several core skills and responsibilities, including problem-solving, resource use, partnering with health care providers, decision-making, self-tailoring, and action planning [7]. These core skills help people with cancer and their caregivers to actively manage the illness and treatment effects, thereby reducing the effects on daily functioning and improving health [7]. For those with multimorbidity, there is more to manage, and engaging in self-management may be more challenging. Interventions that support this complexity are needed.

Digital health tools present opportunities for self-management support for older adults and caregivers [8]. A recent study by Leigh and colleagues [9] found that 68% of older adults aged ≥60 years owned and used smartphones and were interested in using mobile health apps to self-manage heart failure conditions. Mobile health apps have features that can address older adults’ needs and expectations, contributing to enhanced cancer self-management [10-12]. Cancer apps designed for older adults have been reported to enable better communication [13], the potential for patient-reported outcome collection [14], and the feasibility of electronic rapid frailty screening [15-17]. If self- and symptom-management support for older adults with cancer is tailored to the usability and capability preferences of older adults, they can optimize the management of cancer symptoms [7,8]. However, a limited number of mobile health interventions target older adults with cancer and other conditions to support self- and symptom management [10].

In our prior work with older adults, they described both the complex work of cancer self-management and a dearth of supports to manage the complex interplay of their cancer diagnosis with other illnesses [18,19]. Therefore, the purpose of this study was to address this critical gap, by reporting on the user evaluations of the design of a self-management app to support older adults living with cancer and multimorbidity. In this paper, we report on the process and outcomes of this iterative co-design process.

Methods
Study Design
We used the Design Thinking model [20,21], involving a user-centered approach, with engagement from patient partners as coresearchers. A Design Thinking model involves iterative rounds of developing empathy for users, defining functional and usability requirement priorities, and ideating collective perspectives to produce a final prototype [20]. We gathered participant feedback in multiple waves to ensure that the design was user-validated at every step. Figure 1 shows an overview of our methods.
Ethical Considerations

This study underwent ethical review by the harmonized research ethics review board at the University of British Columbia (BREB H21-03052). All participants completed an informed consent process and participated voluntarily. All personal data has been anonymized. All participants were provided an honorarium in the form of a gift card of their choice of $40 per design session.

Recruitment and Participants

Data collection for this study took place between 2021 and 2023 in British Columbia, Canada. We used purposive sampling to recruit older adults with cancer and caregivers. The inclusion criteria for the older adults were aged ≥ 65 years; experiencing or had previously experienced cancer; had received cancer treatment within the previous year; and had at least 1 chronic illness in addition to the cancer diagnosis that required the use of medication or treatment. They also needed to have access and be able to use an internet-connected device, such as a tablet, phone, or laptop. Participants were recruited from community centers, community meetings, support groups, and a registry of patients who had participated in prior research. Caregivers were people who identified as individuals who cared for or supported an older adult during their cancer diagnosis or treatment.

We collected demographic and medical information, health literacy (using the eHealth literacy scale [eHEALS] [22]), and fitness or frailty (using the Vulnerable Elders Survey [23]) scores on all participants, over the telephone, or in person.

We recruited a total of 18 participants comprising 15 older adults and 3 caregivers. In total, 2 older adults participated in evaluating both the low-fidelity and medium-fidelity prototypes to determine whether they perceived any improvements in the design. Thus, a total of 10 participants (8 older adults and 2 caregivers) contributed to the design of the low-fidelity prototype and 10 participants (9 older adults and 1 caregiver) evaluated the medium-fidelity prototype.

Of the 18 participants, 12 (67%) of them were women and 6 (33%) were men. The ages ranged from 40 to 88 years (only age ranges were collected; half [n=9, 50%] of the participants were between 70 and 75 years old). Most were currently married or partnered (n=7, 39%) and college or university educated (n=18, 100%), and they experienced diverse cancers. When asked, they all reported feeling comfortable using the internet.

Defining and Ideating

Using findings from previous work [18] that encapsulated the empathizing stage, we proceeded to define and ideate concepts that could help ground the prototyping phase. We broadly defined this study’s problem statement as supporting older adults’ self-management of their health to improve their quality of life. Using this definition as a prompt, as well as personas to promote empathy with the end users, the research team conducted a brainwriting exercise [24] with the target group to obtain a preliminary understanding of the key tasks that the app should support. A brainwriting exercise is an idea-generation method in human-computer interaction designed to brainstorm and generate what might be a good idea for systems design [24]. From this exercise, we shortlisted tasks that could promote users’ self-management skills according to Grey’s revised self-management theory [25] (Textbox 1). Grey’s revised self-management theory proposes that self-management for both the individual and family should be characterized as interacting with and upon a variety of proximal and distal outcomes [25]. We also highlighted tasks that promoted a more holistic and subjective understanding of users’ health, as those types of tasks could help users think about their quality of life more explicitly.

<table>
<thead>
<tr>
<th>Make daily health reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Report symptoms: a daily report of the user’s symptoms, compared to the day before.</td>
</tr>
<tr>
<td>• Report events: a daily report of events that can impact the user’s symptoms, whether they are physical (taking a walk), emotional (visiting friends), or miscellaneous (weather).</td>
</tr>
<tr>
<td>• Report the day: a daily report of the day that can contextualize symptoms and events.</td>
</tr>
<tr>
<td>• Report questions and notes: a daily report of any questions or notes they may have for their health care provider.</td>
</tr>
<tr>
<td>• Report emoji: a daily report of an emoji that best represents the user’s day. Emojis are effective in comprehension and utility when understanding health reports [26].</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Learn weekly health trends</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Read brief weekly summaries: a textual summary of the trends of a previous week.</td>
</tr>
<tr>
<td>• View visualizations of weekly data: a graphical summary that is equivalent to the textual summary. Informatics have been shown to support holistic wellness by helping older adults with decision-making and identifying trends [27].</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Schedule reminders</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Schedule reminders to be notified of important times (medications, health care visits, and other events) [28].</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Share information</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Email trends and daily reports to health care providers and caregivers, or download them for printing.</td>
</tr>
</tbody>
</table>

Following the brainwriting exercises, we developed a conceptual model [24] to hypothesize, at a high level, how all the tasks should fit together (Figure 2). We leveraged the metaphor of a calendar planner that older adults frequently use to record and keep a journal of their health as a starting point for our model. To validate the usefulness of the model, and to ensure we were on the right track before prototyping, we evaluated the model with 2 participants (P1 and P2) through informal interviews held on Zoom (Zoom Video Communications, Qumu Corporation). As the model was still highly conceptual, their feedback was generally well received, but with the caveat that their positive responses may change once they saw a working prototype.
Development of Prototypes

We first acquainted ourselves with the basic human-computer interaction recommendations for older adults in interface design: minimal new concepts, plain language, unambiguous icons, accessible user interface, larger font and buttons, and consistent visual cues [29]. We also considered how to split information into smaller and more logical steps for more actionable tasks with less cognitive load [30]. This is especially important, as older adults often require more time to learn new computer skills, make more errors, and (generally) need more assistance than younger people [31,32].

We developed a low-fidelity prototype of a minimum viable product using Axure (Axure Software Solutions) and Sketch (Bohemian Coding Company) as our prototyping tools [33,34]. To rapidly assess where the app could be improved (and because we started this work in 2021 at a time when there were still limitations around social gatherings), we remotely evaluated the prototype on Zoom with 8 different participants. Participants showed how they would complete tasks on a browser and verbalized their thoughts, primarily to validate the usefulness of features and assess their high-level usability. We used the Rapid Iterative Testing and Evaluation (RITE) method, which is effective and efficient in identifying and fixing problems [35].

The main problem we identified was that the low-fidelity prototype needed to give more instructions and use plainer language. We also validated whether the visualizations would be easy to interpret, by having participants evaluate several versions of them to assess understandability: those who described having prior training to read graphs (6/8, 75%) highly preferred the visualizations, whereas the others preferred the...
textual summaries. This suggested that it would benefit users to implement both versions (text and graphs) in the app. Overall, the prototype was well received by all 8 participants.

Next, we proceeded with the development of the medium-fidelity prototype using Figma (Figma, Inc) as our prototyping tool [36] (Figure S1 in Multimedia Appendix 1). We also adopted the name Mantra, which draws on the terms “Managing Cancer” and “Comorbidity in Older Adults.” For the medium-fidelity prototype, we focused on the higher fidelity of both aesthetics and interaction compared to the low-fidelity version. The information architecture was determined by the high-level tasks, with more important task flows (ie, daily reports and weekly summaries) explicitly shown in the bottom navigation. Embedded in these 2 task flows was the ability to support information sharing. Scheduling of reminders was placed in the “More” menu (Figure S2 in Multimedia Appendix 1). The “More” menu shows the extent of design capabilities and contains unimplemented tasks to be considered for future iterations.

The task for making daily reports was designed to flow like a web-based questionnaire, with the subtasks taking up steps 1-5 (Figure S2 in Multimedia Appendix 1, shows a symptom report). To alleviate the burden of making daily reports, each step could be skipped and the system autosaved progress so that the user could continue where they had left off. The task for learning weekly health trends was designed so that the textual summary and the visualization for the same week were on the same page but in different tabs to show that they were equivalent (Figure S2 in Multimedia Appendix 1).

Evaluations

We designed the evaluation study to be conducted both remotely and in person, as in-person evaluation may have been too physically demanding for some participants. However, we do acknowledge that the validity of the evaluation may be compromised, as the remote participants were interacting with the prototype through a browser and not on a smartphone, which has different usability concerns. We recruited 6 older adults for remote evaluations and 4 older adults for in-person evaluations of the medium-fidelity prototype.

We had three goals for the evaluation: (1) to understand whether the participants could complete the tasks (assessed through task completion rates); (2) to understand the usability of the prototype (assessed through the System Usability Scale [SUS] [37]); and (3) to understand how the app could be integrated into participants’ existing health management practices (assessed through qualitative interviews after the evaluation of the prototype; see Multimedia Appendix 2 for sample interview questions). The data collected were both qualitative and quantitative. We audio recorded and transcribed all evaluation interviews, which on average lasted 45 (SD 3.0) minutes. Thematic analysis was completed by one of the lead authors with support from the first and final authors [38,39]. We followed the 6-stage approach, which included familiarization, generating initial codes, and searching for themes. The remaining steps related to the naming and thematic structure were refined through an iterative process with weekly meetings between the 2 first authors and the senior author.

Results

Demographic Characteristics of Participants

In total, 15 older adults and 3 caregivers (n=18) participated in this study: 10 participated (8 older adults and 2 caregivers) in the design of the low-fidelity prototype, and 10 evaluated (9 older adults and 1 caregiver) the medium-fidelity prototype (2 older adults participated in both phases). Participants emphasized the importance of tracking functions to make sense of information across physical symptoms and psychosocial aspects symptoms; a clear display; and the organization of notes and reminders to communicate with care providers. The majority of participants were women (8/10, 80%), lived alone (6/10, 60%), and lived at home (10/10, 100%; see Table 1). Only 1 participant was considered frail according to the Vulnerable Elders Survey–13. Most (8/10, 80%) had adequate eHealth literacy according to the eHEALS (mean scores 30.6, SD 9.0), with scores ranging from 8 to 40.
Table 1. Sociodemographic information for all participants: older adults (n=15) and caregivers (n=3).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Participants (N=18), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age group (y)</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;70</td>
<td>4 (22)</td>
</tr>
<tr>
<td>70-75</td>
<td>9 (50)</td>
</tr>
<tr>
<td>76-80</td>
<td>4 (22)</td>
</tr>
<tr>
<td>81-85</td>
<td>0 (0)</td>
</tr>
<tr>
<td>&gt;86</td>
<td>1 (6)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>6 (33)</td>
</tr>
<tr>
<td>Women</td>
<td>12 (67)</td>
</tr>
<tr>
<td><strong>First language</strong></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>18 (100)</td>
</tr>
<tr>
<td><strong>Housing situation</strong></td>
<td></td>
</tr>
<tr>
<td>At home (house, condo, or apartment)</td>
<td>18 (100)</td>
</tr>
<tr>
<td><strong>Living situation</strong></td>
<td></td>
</tr>
<tr>
<td>Alone</td>
<td>9 (50)</td>
</tr>
<tr>
<td>Spouse</td>
<td>8 (44)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (6)</td>
</tr>
<tr>
<td><strong>Marital status</strong></td>
<td></td>
</tr>
<tr>
<td>Married or living common law</td>
<td>7 (39)</td>
</tr>
<tr>
<td>Widow or widower</td>
<td>3 (17)</td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>5 (28)</td>
</tr>
<tr>
<td>Single (never married)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (5)</td>
</tr>
<tr>
<td><strong>Education level</strong></td>
<td></td>
</tr>
<tr>
<td>In total, 13 years and more (some or completed college or university)</td>
<td>18 (100)</td>
</tr>
<tr>
<td><strong>Type of cancer treatment (could select more than 1 type; n=15)\textsuperscript{a,b}</strong></td>
<td></td>
</tr>
<tr>
<td>Surgery</td>
<td>13 (87)</td>
</tr>
<tr>
<td>Radiation</td>
<td>6 (40)</td>
</tr>
<tr>
<td>Chemotherapy, targeted therapy, or immunotherapy</td>
<td>10 (67)</td>
</tr>
<tr>
<td>Hormone therapy</td>
<td>5 (33)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (7)</td>
</tr>
<tr>
<td><strong>Treatment intent as reported by patient</strong></td>
<td></td>
</tr>
<tr>
<td>Curative</td>
<td>14 (93)</td>
</tr>
<tr>
<td>Palliative</td>
<td>2 (13)</td>
</tr>
<tr>
<td><strong>Current comorbidities</strong></td>
<td></td>
</tr>
<tr>
<td>Asthma, emphysema, chronic bronchitis, or COPD\textsuperscript{c}</td>
<td>4 (22)</td>
</tr>
<tr>
<td>Arthritis or rheumatism</td>
<td>12 (67)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Digestive problems (ulcer, colitis, and gallbladder disease)</td>
<td>5 (28)</td>
</tr>
<tr>
<td>Heart trouble (angina, congestive heart failure, or coronary artery disease)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Depression or anxiety</td>
<td>2 (11)</td>
</tr>
</tbody>
</table>
Task Completion and Usability

In terms of task completion, we found that 90% (9/10) of participants who evaluated the medium-fidelity prototype were able to complete all of the tasks as scored and presented in Multimedia Appendix 3. However, 1 participant (P18) experienced initial difficulty reading the text on the screen and would not have been able to complete the “Report Symptom” feature if not for the hints we provided. Nevertheless, this high completion rate indicates that the app was sufficiently designed for older adults in our target population to accomplish the key tasks.

We also assessed the general usability of the app, by administering the SUS, and found that participants evaluated the system as very usable, with an average of 87 which is described as the “best imaginable” according to Bangor and colleagues [37]. Table 2 shows the individual scores for each participant, broken down by question. Furthermore, according to Sauro and Lewis [40], the scale can be defined by 2 dimensions, that is, learnability (questions 4 and 10) and usability (all other questions). As seen in Table 2, only P18 struggled with learnability while all others perceived the app to be easy to learn. In sum, the high SUS score from the user evaluations, reflecting high user satisfaction and usability, shows that this app has great potential to assist older adults in their self-management activities.

Table 2. System Usability Scale responses.

<table>
<thead>
<tr>
<th>Participant</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>P3</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>92.5</td>
</tr>
<tr>
<td>P9</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>P11</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>80</td>
</tr>
<tr>
<td>P12</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>95</td>
</tr>
<tr>
<td>P13</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>90</td>
</tr>
<tr>
<td>P14</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>97.5</td>
</tr>
<tr>
<td>P15</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>97.5</td>
</tr>
<tr>
<td>P16</td>
<td>3</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>70</td>
</tr>
<tr>
<td>P17</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>5</td>
<td>1</td>
<td>92.5</td>
</tr>
<tr>
<td>P18</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>55</td>
</tr>
</tbody>
</table>

Table explanation: there are 10 questions in total, ranging from strongly disagree (1 point) to strongly agree (5 points). The tone of the questions (see Multimedia Appendix 3) switches from positive (odd questions) to negative (even questions). The score for each person is calculated as follows: \[ X = \text{sum of the points for all odd-numbered questions} - 5; \ Y = 25 - \text{sum of the points for all even-numbered questions}; \text{individual SUS score} = \left( X + Y \right) \times 2.5. \] An average score of 85-100 is the best imaginable, while 52-85 is still considered excellent [37]. All participants evaluated the system as having high usability, except P18 who strongly agreed with the need for support from a technical person or additional knowledge to comfortably use the app.

Qualitative Findings

Overview

Based on our thematic analysis, we constructed 3 global themes from the data, related to both the nature of the app and its value in supporting self-management. These themes were organized and labeled as follows: (1) app conceptual model matches users’ mental model, (2) value and usefulness for self-management, and (3) confusing icons and buttons. These themes are discussed further below.

App Conceptual Model Matches Users’ Mental Model

Overview

In our user testing, participants found the app interfaces easy to navigate with straightforward steps to complete tasks. They also found the text summaries and data visualizations to be visually engaging with clear health-related feedback, which contributed to the ease of use of the app. In light of these user experiences, the following subthemes were constructed from the data, related to the app structure, its usability, and its ease of use: (1) transferability; (2) intuitiveness and ease of navigation; (3) clear data visualization; and (4) simplified and comprehensive summary reports.
Transferability

In our user testing, participants found the app to have a familiar interface, with similar features to other popular health apps that they had used or were using, which reinforced a sense of transferability. For example, 1 participant stated: “I would say that the labels were very clear to me which are similar to some of the apps I’ve used” (P15). Participants also described the app prompts as predictable with simple steps to complete tasks and found the app’s phrases and concepts to be familiar and understandable. Overall, the words and concepts used in the app followed real-world conventions which made the information appear in a logical order for participants. One of the participants said: “It was very predictable and understandable using it the first time” (P12).

Intuitiveness and Ease of Navigation

The app user interface offered a sharp, constant, and uncluttered background, which enhanced readability. Participants felt that the steps and buttons needed to complete the tasks were easy to navigate, which motivated them to use the app. They appreciated the simplicity of the app features which contributed to its intuitiveness. One participant commented: “I think even an elementary student could probably use this app quite easily” (P12). Furthermore, the simple interface and the easy task flow from the home page to completion, shaped user experiences as reflected here: “It was easy to use...the app is simple...but these other apps have a lot and it’s always difficult” (P14).

For some participants, the app features felt orderly and streamlined, increasing accessibility even for users with color sensitivity. The legibility and high-contrast colors of the interface alongside the large font size were perceived as being visually engaging and readable. Some participants suggested an adjustable font size should be embedded in the app to allow customization: “Everything looks clear except that the fonts needed to be adjusted big enough. I was straining to see some things” (P9).

Overall, participants found the app easy to use regardless of their level of experience with technology and cancer type. Although older adults reported feeling comfortable using the app; some users required guidance to navigate certain icons on the app. Specifically, the remote participants had more difficulty navigating the prototype if they used a phone remotely than in-person participants. One participant stated: “I had to ask for help to identify icons on the phone” (P11). Another said: “You know, I could go through it [the app] but I needed some help” (P9).

Clear Data Visualization

Participants’ most desired accurate unambiguous feedback with simple and clear data visualizations. Participants perceived the clarity and accuracy of the data visualizations as visually appealing, engaging, and readable.

While acknowledging color contrast in the charts and graphs, participants described how texts and labels highlighted important information. Participants reported that they could easily understand the visual elements, facilitating the effective abstraction of actionable insights from the data visuals. For instance, 1 participant stated: “The texts are fine with me...I can understand the data from the graphs” (P11). However, 1 participant highlighted their difficulty understanding the meanings of the different colors included in the charts. As a remedy, they suggested 1 color stream to ensure consistency. They also suggested that emojis should be used in place of graphs since emojis are considered a universal language for most users. Additionally, the text or data presented in the graphs might be difficult to understand and interpret by some non-English users of the app. As presented to participants, the app allowed for only 1 representative emoji per day. However, users suggested combining many emojis that best represent their emotions and feelings for the day. For example, users could combine 2 or more emojis to illustrate their emotions since they could have different experiences within a day—positive, neutral, and negative emotions: “I prefer more emojis, in fact a combination of them...because they tell more story about my emotions” (P9). Participants noted clear data visualization and accuracy of information displayed and affirmed the relevance of having readable and understandable trends of their health data tracked and presented to them in graphs covering either a month or week period.

Simplified and Comprehensive Summary Reports

Participants observed that the clear and comprehensive nature of the health data summaries helped to make connections between day-to-day emotions and changes over time. Some users noted the easy, actionable steps of reporting symptoms and feelings that were directly linked to the feedback. Participants found that displaying multiple types of data on 1 page was more intuitive and comprehensive and helped them make sense of their overall health:

Usually, what I see when I look at my blood reports [in leukaemia] is, I click on a report and I see a single trend line [in another app]. Now, with this app, I can see all different types of reports on one page. That’s good. It keeps it together. Because I can see that it’s helping me make connections of all the data. [P15]

In comparing this app with other apps, participants noted that this app tends to present an integrated analysis of their health, emotions, and present health state, which made it more interesting to track and review. For example, 1 older adult stated:

The Apple Health app was the only one that I have used. And that tends to be reflecting sort of like a simplified quantitative analysis, not qualitative, not including emotions and things. And I think that would be an interesting thing to record and then to review. The correlation between my actual situation of cancer and how I felt about it at the time. [P15]

Participants appreciated summary reports, although some older adults expressed a preference for monthly summary reports rather than weekly reports. They argued that cancer is a long-term condition that requires a summary of the symptoms over a longer period. Further, 1 participant said: “I would like a monthly report summary...the other apps I do it monthly...I will go over and look at the trend lines over the months. This would fit into that whole world that I’m a part of and have been now since 2015” (P3).
Value and Usefulness for Self-Management

Overview
Participants described how the app was valuable and useful in supporting them to manage their health, in the presence of a cancer diagnosis. Participants felt that the various dimensions of the app could support self-management as captured and explained in the following subthemes: (1) integration into existing self-management routines; (2) app design linked to user personal health needs; (3) recording and sharing of health information; and (4) digital health calendar.

Integration Into Existing Self-Management Routines
Our user testing demonstrated that users were ready and willing to use the app for self-management and felt the app could help them meet their health needs. App features such as the daily symptom report—which required users to report symptoms daily at a time convenient to them—were perceived as a normal daily activity that could easily be done without hassle. Participants found the app’s notifications feature a useful way to engage them and remind them to report their feelings, adhere to treatments, and honor appointments with their clinicians. For example, a participant observed that the notifications could easily be synchronized with other daily tasks:

I could see myself spending probably half an hour a day inputting information and then reviewing the results. [With another app], every morning, I wake up and check what my sleep score was and what my resting heart rate was. And when my resting heart rate goes up, it tells me that I’m not behaving myself, or stressed. So, it’s just kind of a check-in with myself almost to make sure that I’m, you know, kind of following the health path that I want to be on. Again, the reminder notification is great for reporting my emotions. [P15]

To encourage consistent and continuous engagement of the app in daily self-management, the app was designed to prompt users to input their health report daily.

Participants expressed that reporting their symptoms and interventions daily in the app would not impact other life activities or routines and they never found tracking in the app troubling or burdensome. Further, 1 participant said:

For me it would be because I don’t have many other obligations. I’m retired. And my wife and I live at home. We don’t travel much. Not at all really. So, I mean to us, you know, this is the sort of thing that I do every day. I’m on the computer doing stuff like this [P11]

However, some participants felt there might be user resistance during the initial stage of the app implementation particularly for those who are not technologically savvy and may also not have a family member to guide them to navigate the app. Further, 1 participant speculated about possible resistance: “I think there’s going to be some resistance to adapt for those who are not good at the computer. It’s doable, but it’s going to be difficult for them” (P3). Overall, participants expressed high intentions of integrating the app into their daily routines for self-management.

Design Linked to User Personal Health Needs
Participants reported that the app design supported their individual health needs—symptom-monitoring, tracking, and self-managing. While the app was perceived to keep users motivated, they also felt it was designed with user needs at the forefront. Further, 1 participant, who had an adult son with autism who needed to track his diet, felt the app could be useful not only in monitoring their own treatment-related symptoms, but also for noncancer patients: “Excellent, this app could align well to my health needs and also probably work well with my son too” (P3). Older adult users could in real time monitor their unique cancer symptoms, track medication adherence, and receive reminders for appointments and other relevant health updates. Further, 1 participant stated:

Well, I mean I wanted to set some goals in terms of, you know, trying to fit exercise in, and it was a way for me to challenge myself. I guess, the features that I do use are here [on this app] to receive notifications and track my health. [P9]

Recording and Sharing of Health Information
Participants reported the value of the app’s capability to record and track information for later sharing with their clinicians. This act of keeping track of information in the form of note-taking was perceived as a better alternative to electronically record, share, and communicate users’ health issues with their health care providers than other analogue strategies. Further, 1 older adult said:

It’s a better way to organize notes and trends and see about trends. Just as a communication tool...maybe a better communication tool to use to talk to the different healthcare providers that we communicate with...rather than recording in so many places. [P13]

Some participants observed that the app features mimic what older adults usually do in keeping track of events by documenting on a piece of paper their daily feelings, thoughts, and other related health issues. Some participants felt that better symptom tracking helped to gain control over symptoms and improve general well-being. Further, 1 participant stated:

So what this does is that you know. So, this feature imitates older adults, what they do with their own, like say journal or calendar or anything that they write things on. They keep track of their questions or note what they’re feeling. [P12]

Older adults repeatedly voiced their willingness to use the app because it offered them the ability to share their health information with their clinicians during appointment visits. Most participants felt that clinicians could deliver more tailored care when provided with additional information during consultations.

Digital Health Calendar
All necessary details regarding patients’ symptom-tracking reports, treatments, and cancer education, including schedules and reminders or alerts are diarized for future reference.
Participants reported how app features particularly the “reporting” and “summaries” features provide a safe place to store more secure information in the app for reference. Further, 1 participant said:

*It’s like an electronic diary...a good calendar to write and keep track of notes about their feelings...So you have that instant data that you can look at where you’ve been and where you’re going.* [P14]

In addition to storing information related to a person’s cancer, the app also makes it easier for the user to edit and update information with new tasks and reschedule appointments. According to participants, the app guides users to know their current state of health and to predict their future health by simply referring to the calendar. For instance, 1 older adult stated: “I mean, one of the things you can do is to keep information here [in app]...you can schedule new appointments with your doctor. It’s like I could see it having lots of benefits” [P11].

Other participants emphasized the value and usefulness of the app for older adults in keeping their health records. Participants argued that most older adults are forgetful as a consequence of aging, and this app could be a great tool to help them remember their changing health. Further, 1 participant stated:

*I mean, one of the things is as we age, if your memory is having trouble with remembering things, that probably helps...it would help that because then you could go back and when was I really feeling so bad and you know? That’s great to know.* [P9]

For some participants, this app could also help in tracking different types of treatment and associated side effects, especially for users starting new treatments. Participants also emphasized the value of this tool early in the treatment process, given the capability to track changes over time and remind users of appointments to and schedules. A participant stated:

*Well, what I do is I have to go through and scroll every appointment. But where this is useful for me is when, especially when I’m starting a new treatment that I’ve never done before. I want to track the side effects and I want to know what the trigger points are on it...the notification reminders, it’s like a reminder of things that you have to experience. Those kinds of elements that I want to be able to look at and see how that changed over six months.* [P12]

**Confusing Icons and Buttons**

The last theme relates to opportunities to improve the app based on evaluations. Despite the intuitiveness and ease of navigation of the prototype, some participants found the icons and buttons confusing while others could not understand some features and information displayed. These attributes impeded usability. Further, 1 participant noted:

*Consider seeking medical care. Is this with my GP? Is this with the oncologist? Do I call for medical care for fatigue? Community health news, what does this button mean? I also wonder what that recommendation is.* [P3]

Another participant expressed their dissatisfaction with not being able to find suitable emojis to express their emotions and asked: “I don’t even see happy on the list here. Where can I find happy emoji? It’s rather lots of other things here” [P9]. While this participant could not find suitable emojis to express their emotions, others too found the buttons too hard to tap causing frustration as they engaged with the app prototype. A participant stated:

*...because I’m not fluent with how the technology always is depicted. I tend to tap, tap, and click, click, click fast. It’s hard...It draws, but has been crazy and often I freeze something because I’ve tried to make it all go too quickly right and I’ve not given the computer time to catch what I’m doing.* [P13]

Participants also found some icons not matching their current state. For example, participants who are used to seeing a little pencil icon indicating “write something more” were confused when it was rather meant for “edit.” A participant stated:

*I'm accustomed to seeing...the little pencil with respect to writing something more. So I think I would have assumed I should write something about nausea as opposed to edit my answer...And it’s possible that people don’t realize [it].* [P3]

**Discussion**

**Principal Findings**

In this paper, we present the findings of an iterative co-design and evaluation study of a cancer self-management app prototype designed specifically for the needs of older adults living with cancer and multimorbidity. Our key finding was that creating a space for this population to track and interpret data related to their health, in a way that made sense to them, would support their self-management needs. The app we designed built on the concept of the calendar, a routine activity that many older adults already participate in. By drawing on this common routine we developed an app that is both useful and acceptable and was not considered burdensome. This work is novel as it acknowledges the complex health states that this population experiences and aims to address their challenges through an app that was co-designed with them, for them. Our approach was user-centered, with a research team co-led by older adults with cancer experiences and comorbidities.

**Limitations**

The main limitations of this study relate to the sampling. First, most of the participants were not on active treatment at the time of study participation. Future studies of the app should study usability for those with a new diagnosis using the app for the first time. Second, most participants spoke English, making usability generally easier. Third, all participants were college or university educated. Future studies should include participants with secondary education or lower to gather more broader and diverse perspectives across all levels of education and experience for the design. Finally, many of the participants regularly used technology and smartphones—however, 2 scored low on the eHEALS measure of health literacy. Nevertheless, we observe growing rates of smartphone use among older adults.
Comparison With Prior Work

This study underscores older adults’ interest in proactively managing their own cancer diagnosis and other existing conditions. While older adults’ interest in self-management has been reported in other studies [5,43], our study highlights the functionalities that support this work. Given that many older adults are managing distressing symptoms as a result of cancer and its treatments, alongside existing chronic illnesses, the impetus to mobilize self-management supports is critical [44]. Our app in its current state is designed for all older adults with cancer and other illnesses to self-manage their conditions, but it is most useful and valuable for older adults who have just been diagnosed with cancer.

Through our iterative co-design process, several noteworthy features were emphasized by older adults involved in this study. One of the features most valued by participants was the ability to facilitate meaningful connections across all aspects of health. A recent scoping review by Wilson and colleagues [45] found that apps that do not support older adults to make connections across functions were a key barrier to meaningful engagement with apps. The app developed in this study provides direction and guidance on how to report symptoms, events, and daily emotional states and provides an easy display to integrate these various inputs and share insights with health care providers. Unlike existing mobile apps with several views on multiple interfaces [46-48], this app provides a single view of all symptoms, events, and emotions consolidated on 1 page which aims to improve usability, efficiency, and satisfaction among users. The design also allows an effective presentation of daily reports in a comprehensive and easy-to-understand manner. Additionally, the app’s design exhibits flexibility, allowing older adults—in collaboration with their health clinicians and caregivers—to prioritize specific symptoms (or variables) of focus. These features are in sharp contrast to existing apps and websites used by older adults to track health information, which often have a narrow scope, concentrating solely on particular cancers or symptoms [27,49-52]. Health apps with narrow scopes may hamper a comprehensive understanding of health and quality of life from a multidimensional perspective [27].

While other studies detail older adults reporting difficulties with limited functionalities and comprehension of visual health data [45,53], participants in this study emphasized a preference for both text and visual summaries to help them make sense of patterns in the data. Older adults also appreciated and valued different modes to view data trends. For example, notes and textual summaries served as an alternative means of conveying information for those who struggled with graphical representations. This has been reported in other studies both within cancer [54] and elsewhere [55]. For example, in a recent pilot randomized controlled trial by Lally and colleagues [54], older adults expressed a preference for information leaflets and text notes to describe the trends of their health. Our study reiterates those findings and also highlights the possibility of text summaries becoming “a voice” for older adults during consultations with their clinicians. Text summaries allow for better and more streamlined conversations with clinicians, thereby promoting shared decision-making [56,57]. It is important to note that the preferences and recommendations of our participants informed revisions of the app features and functionalities in the prototype iterations, and the current state of the app reflects the needs of its potential users. Overall, the app developed in this study represents an acceptable and usable app that is adaptable to the unique needs of older adults with cancer in monitoring changes in their health.

Conclusions

Our self-management app prototype has both content and face validity among older adults with cancer and comorbidities. At this stage, the app requires further refinements and testing to understand its efficacy and to gauge its acceptability and implementation potential within the cancer care system in Canada and beyond.

Acknowledgments

The authors wish to acknowledge the funding provided by the Oncology Nursing Foundation (RE03) and a Scholar award from the Michael Smith Foundation for Health Research BC (Kristen R Haase). Dr Puts is supported by a Canada Research Chair in the care of frail older adults.

The authors would like to gratefully acknowledge the significant, methodological, and scholarly contributions of Margaret Tompson.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Medium-fidelity prototype.
[DOCX File, 1682 KB - aging_v711e53163_app1.docx]
Multimedia Appendix 2
Interview sample questions.

[DOCX File .15 KB - aging_v7i1e53163_app2.docx]

Multimedia Appendix 3
System Usability Scale questions (rated between strongly agree to strongly disagree).

[DOCX File .14 KB - aging_v7i1e53163_app3.docx]

References


https://aging.jmir.org/2024/1/e53163

JMIR Aging 2024 | vol. 7 | e53163 | p.488
(page number not for citation purposes)


Abbreviations

- eHEALS: eHealth literacy scale
- RITE: Rapid Iterative Testing and Evaluation
- SUS: System Usability Scale
Remote Evidence-Based Programs for Health Promotion to Support Older Adults During the COVID-19 Pandemic and Beyond: Mixed Methods Outcome Evaluation

Lesley Steinman¹, MPH, MSW, PhD; Kelly Chadwick¹, MPH; Erica Chavez Santos¹, MPH, PhD; Sruthi Sravanam¹, BTech, MS; Selisha Snowy Johnson², MPH; Elspeth Rensema³, MPH; Caitlin Mayotte¹, MS; Paige Denison⁴, BS; Kate Lorig⁵, DrPH

¹Health Promotion Research Center, Department of Health Systems and Population Health, University of Washington School of Public Health, Seattle, WA, United States
²Office of Community Outreach and Engagement, Fred Hutch Cancer Center, Seattle, WA, United States
³Department of Anesthesiology and Pain Medicine, University of Washington School of Medicine, Seattle, WA, United States
⁴Sound Generations, Seattle, WA, United States
⁵Self Management Resource Center, Aptos, CA, United States

Abstract

Background: Evidence-based programs (EBPs) for health promotion were developed to reach older adults where they live, work, pray, and play. When the COVID-19 pandemic placed a disproportionate burden on older adults living with chronic conditions and the community-based organizations that support them, these in-person programs shifted to remote delivery. While EBPs have demonstrated effectiveness when delivered in person, less is known about outcomes when delivered remotely.

Objective: This study evaluated changes in remote EBP participants’ health and well-being in a national mixed methods outcome evaluation in January 1, 2021, to March 31, 2022.

Methods: We used the RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) for equity framework to guide the evaluation. We purposively sampled for diverse remote EBP delivery modes and delivery organizations, staff, and traditionally underserved older adults, including people of color and rural dwellers. We included 5 EBPs for self-management, falls prevention, and physical activity: videoconferencing (Chronic Disease Self-Management Program, Diabetes Self-Management Program, and EnhanceFitness), telephone plus mailed materials (Chronic Pain Self-Management Program), and enhanced self-directed mailed materials (Walk With Ease). Participant and provider data included validated surveys, in-depth interviews, and open-ended survey questions. We used descriptive statistics to characterize the sample and the magnitude of change and paired t tests (2-tailed) and the Fisher exact test to test for change in outcomes between enrollment and 6-month follow-up. Thematic analysis was used to identify similarities and differences in outcomes within and across programs. Joint display tables facilitated the integration of quantitative and qualitative findings.

Results: A total of 586 older adults, 198 providers, and 37 organizations providing EBPs participated in the evaluation. Of the 586 older adults, 289 (49.3%) provided follow-up outcome data. The mean age of the EBP participants was 65.4 (SD 12.0) years. Of the 289 EBP participants, 241 (83.4%) were female, 108 (37.3%) were people of color, 113 (39.1%) lived alone, and 99 (34.3%) were experiencing financial hardship. In addition, the participants reported a mean of 2.5 (SD 1.7) chronic conditions. Overall, the remote EBP participants showed statistically significant improvements in health, energy, sleep quality, loneliness, depressive symptoms, and technology anxiety. Qualitatively, participants shared improvements in knowledge, attitudes, and skills for healthier living; reduced their social isolation and loneliness; and gained better access to programs. Three-fourths of the providers (149/198, 75.2%) felt that effectiveness was maintained when switching from in-person to remote delivery.
Conclusions: The findings suggest that participating in remote EBPs can improve health, social, and technological outcomes of interest for older adults and providers, with benefits extending to policy makers. Future policy and practice can better support remote EBP delivery as one model for health promotion, improving access for all older adults.

(JMIR Aging 2024;7:e52069) doi:10.2196/52069

KEYWORDS
older adults; health equity; rural; chronic disease; outcome evaluation; behavior change; technology; community based; evidence based; health promotion; mobile phone

Introduction

Background

One in 6 adults living in the United States is an older adult (aged ≥60 y). This number is expected to double in the next 40 years [1]. While older adults possess the wisdom of experience and are often actively involved in taking care of themselves, their family members, and their communities, they also face a wide range of unique health challenges that come with aging [2]. A disproportionate number of older adults live with chronic health conditions: 85% have 1 chronic health condition, and 60% report managing at least two [3]. One in 4 older adults reports at least 1 fall every year, and falls remain a leading cause of death and injury among this age group [4]. These health issues—combined with social determinants of health such as older adults’ built environment, social context, and access to medical care—put older adults at risk for premature death and poorer quality of life [5]. While only accounting for 17% of the population, older adults make up 35% of health care costs, according to 2019 Medical Expenditure Panel data [6]. Furthermore, 86 cents of every dollar of health care spending goes toward chronic conditions [7,8], and the burden of chronic disease is unevenly borne by women; older adults; people of color; and people living in poverty who experience disparities in access to, and quality of, care [9].

To address these challenges, many health promotion programs have been created based on the Chronic Care Model [10], to improve access and quality of care through community-based programs that teach knowledge, skills, and self-efficacy to enhance older adults’ health and well-being in their daily lives. Among these are evidence-based programs (EBPs) that have been researched and recognized by national and federal agencies, such as the Administration for Community Living (ACL) and the Centers for Disease Control and Prevention (CDC) Arthritis Management and Well-Being Program, as being effective in promoting health outcomes through standardized interventions [11,12].

Investigating the Effectiveness of Remote EBPs

Before the COVID-19 pandemic, many of these EBPs were primarily offered in person. This was not possible during the pandemic due to safety guidelines regarding physical distancing. At the same time, there was an increased need for programs because older adults became more isolated and less physically active, and they had less access to an overtaxed medical system [13,14]. Seeing this need, EBPs quickly pivoted to remote delivery by mail, telephone, videoconferencing, or a combination of these modes. While researchers suggest that EBPs will work as intended using different forms as long as core functions are not modified [15,16], it is unknown whether program effectiveness is maintained when switching from in-person to remote delivery. To date, there have been limited studies investigating the effectiveness of remote EBPs that were originally designed to be offered in person [17-21]. As such, we conducted a longitudinal national outcome evaluation of several EBPs to assess changes in older participants’ health.

Methods

Framework and Design

We used the RE-AIM (Reach, Effectiveness, Adoption, Implementation, and Maintenance) for equity framework [22] to evaluate the potential impact of remote EBPs on older adults’ health and well-being. The equity lens means that in addition to evaluating impact, we looked at outcomes across programs, sampled organizations that reach older adults who are underserved, reported EBP reach, and assessed whether there were any unintended consequences. As such, we used mixed methods to “give voice to participants as well as report statistical trends” [23]. We conducted a multisite single-group pre-post evaluation. Multiple sites were selected to facilitate the generalizability of the findings, and a single-group design was chosen to make it feasible to conduct the evaluation during the COVID-19 pandemic and due to the descriptive nature of this study.

Ethical Considerations

This study was considered exempt from University of Washington Institutional Review Board review because the activities fell under category 2 with regard to quality improvement and program evaluation (STUDY00011549). Participants who completed the preprogram survey were given a US $10 gift card, and participants who completed the follow-up survey were given a US $20 gift card (we provided electronic gift cards unless participants requested a physical card). Providers who completed the survey received a US $10 electronic gift card.

Selected EBPs

EBPs are health promotion programs that have been evaluated and proven to be effective. We included 5 EBPs in our evaluation: Chronic Disease Self-Management Program (CDSMP), Diabetes Self-Management Program (DSM), EnhancementFitness (EF), Chronic Pain Self-Management (CPSM), and Walk With Ease (WWE). These EBPs are all currently recognized by the national social services agency for older adults (ACL) and public health agencies (CDC) [11,12]. Before the
COVID-19 pandemic, these programs were offered in person in group formats; they were adapted as follows for remote delivery in response to the COVID-19 pandemic: videoconferencing (CDSMP, DSMP, and EF), telephone plus mailed materials (CPSMP), and self-directed mailed materials plus enhanced support (WWE).

These 5 EBPs were selected because they had sufficient program reach and represented different remote modes (telephone, videoconferencing, mail, or a combination of these modes) and health topics (chronic disease self-management, falls prevention, and physical activity). The adaptation process largely involved planned changes to modify delivery context (eg, smaller class size; adding a cofacilitator to support engagement with, and the usability of, technology; and providing telephone options for people without access to reliable internet or videoconferencing technology) as well as some unplanned changes that occurred organically during field experience in consultation with the program leads (rather than changes to program content) [24]. It should be noted that WWE had created an enhanced self-directed program before the pandemic; we included this program in the evaluation to assess different modes of remote EBP delivery, given partners’, policy makers’, and older adults’ interest in diverse ways to engage in remote health promotion. The enhanced self-directed WWE is delivered via a cohort, with a trained leader providing web-based motivation and support to individual participants during the program period. More information about the included programs is provided below (Figure 1) and on each program’s website, including guidance for remote delivery.

**Figure 1.** Participating remote evidence-based programs.

<table>
<thead>
<tr>
<th>Chronic Disease Self-Management Program</th>
<th>Chronic Pain Self-Management Program</th>
<th>Diabetes Self-Management Program</th>
<th>EnhanceFitness Program</th>
<th>Walk With Ease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focuses on techniques for dealing with symptoms of chronic conditions</td>
<td>Focuses on techniques for dealing with symptoms of chronic pain conditions</td>
<td>Focuses on techniques for managing diabetes</td>
<td>Focuses on exercise to prevent falls</td>
<td>Focuses on physical activity and education for arthritis management</td>
</tr>
<tr>
<td>Group based or individual</td>
<td>Group based or individual</td>
<td>Group based or individual</td>
<td>Group based</td>
<td>Individual</td>
</tr>
<tr>
<td>Offered as 6 weekly 2.5-hour videoconferencing sessions</td>
<td>Offered as a mailed tool kit of materials that with 6 weekly scripted 1-hour small-group telephone calls</td>
<td>Offered as 6 weekly 2.5-hour videoconferencing sessions</td>
<td>Offered as 1-hour sessions, 3 times per week, on an on going basis via videoconferencing sessions, with fitness tests at baseline, at 16 weeks, and at regular intervals thereafter</td>
<td>Offered as a self-directed enhanced program, with a leader coordinating communication to a cohort of participants who separately move through the program over 6-weeks</td>
</tr>
</tbody>
</table>

**Recruitment**

In early 2021, we recruited organizations delivering remote EBPs with a brief web-based interest survey and webinars through several networks: EBP training listserves, the CDC Arthritis Program and ACL grantees, the Evidence-Based Leadership Collaborative, and regional EBP networks. The organizations included social services, public health, and health care agencies that were largely community based, although some were located in clinical settings. We used maximum variation purposive sampling [25] to identify organizations of diverse sizes and types, varied provider characteristics, and from different geographic areas to aid the generalizability of the evaluation findings. This sampling prioritized organizations engaging older populations with low-income status with multiple chronic conditions who are vulnerable to both COVID-19 infection and poor health outcomes and care (eg, people of color, those with disabilities, and those living in rural areas). The eligibility criteria for organizations was the delivery of at least 1 of the 5 remote EBPs from January 2021 through March 2022. We then used convenience sampling to invite all remote EBP participants to take part in the evaluation. Participant surveys were primarily conducted on the web using REDCap (Research Electronic Data Capture; Vanderbilt University) [26], with options to complete surveys by telephone or by mail.
Participants were surveyed when they enrolled in the program and again 6 months after program enrollment (regardless of when they finished the program). At both time points, a link to the survey was emailed to participants, and they received up to 3 additional reminder emails and 3 reminder calls during the month of survey eligibility. In addition, community-based organizations encouraged baseline survey completion as part of program intake and orientation.

Furthermore, all EBP providers at participating organizations were invited to take part in a 1-time survey eliciting their perspectives regarding the impact of remote EBPs on participants as well as providers. The EBP providers included leaders (people who delivered remote programs directly) and managers (people who coordinated and oversaw program delivery).

Data Collection

**Quantitative**

The participant outcome survey (Multimedia Appendix 1) combined several brief self-rated health measures that have been validated with older adults: single-item self-rated health [27,28], pain [29], fatigue or energy [30], loneliness (University of California Los Angeles 3-item Loneliness Scale [31]), social isolation (4-item Social Network Index [32]), depression (Patient Health Questionnaire [PHQ]-8 [33]), anxiety (Generalized Anxiety Disorder-2 [34]), physical activity (Exercise Vital Sign [35]), and social needs [36]. These cross-cutting outcomes were selected in partnership with organizations, developers or administrators, and policy makers to identify key outcomes of interest across the health promotion programs.

We also collected several program-specific outcomes [37] related to their health focus: self-efficacy for CDSMP, hypoglycemia for DSMP, and pain and the use of opioid medications for CPSMP. In addition, the preprogram survey included questions about remote EBP participant demographics: age, gender, race, ethnicity, living alone, caregiving, and chronic conditions. Rurality was defined using the participant’s zip code and several federal criteria for rural funding [38]. The 6-month follow-up survey included 3 additional quantitative items: two examining the usability of and anxiety about technology using the Senior Technology Acceptance Measure [39] and the other (Social Network Index [32]), depression (Patient Health Questionnaire [PHQ]-8 [33]), anxiety (Generalized Anxiety Disorder-2 [34]), physical activity (Exercise Vital Sign [35]), and social needs [36]. These cross-cutting outcomes were selected in partnership with organizations, developers or administrators, and policy makers to identify key outcomes of interest across the health promotion programs.

Provider survey data included multiple-choice questions about demographics, experience delivering EBPs, and the impact of remote EBPs on both participants and providers. Response choices were created from open-ended responses to an earlier Evidence-Based Leadership Collaborative remote EBP web-based survey in 2020 (L. Steinman, personal communication, December 2020).

**Qualitative**

The 6-month follow-up participant survey included open-ended items about program acceptability and unintended consequences or impacts. Participants who completed the follow-up survey over the telephone were also asked 2 additional qualitative questions about participating in the evaluation. Providers were also asked an open-ended question about any additional benefits for both remote EBP participants and providers.

Data Analysis

**Quantitative**

Data were managed in REDCap [26] and analyzed using R software (R Foundation for Statistical Computing) [43]. Our primary aim was to evaluate changes in remote EBP participants’ health and well-being (self-rated health, fatigue or energy, exercise, pain, sleep, depression, anxiety, loneliness, social isolation, and program-specific outcomes) between program enrollment and 6-month follow-up. Only participants who included both baseline and 6-month follow-up data were included in these analyses. We assessed the magnitude of the changes using descriptive statistics, percentage change, and Cohen $d$ effect sizes and tested for statistical significance of the changes using paired $t$ tests (2-tailed). Percentage change was reported for both people who improved—or maintained—outcomes between remote EBP enrollment and 6-month follow-up, given the importance of maintaining health in an aging population [44]. Effect sizes were calculated as mean (SD) [45]. We also used descriptive statistics to characterize the sociodemographic characteristics of remote EBP participants and providers. We chose not to use the Bonferroni correction to adjust the $P$ values because we carried out tests on multiple outcomes of interest across programs without a priori hypotheses (rather than 1 primary outcome of interest) [46].

**Qualitative**

Audio-recorded data were transcribed into Microsoft Word documents. These text data from transcripts and open-ended survey questions were analyzed using Dedoose software (SocioCultural Research Consultants, LLC) [47]. For qualitative data, we used thematic analysis [48,49] to understand similarities and differences across and within remote EBP acceptability and benefits or unintended consequences. Two researchers (KC and LS) created a codebook to organize important text for comparison, using deductive codes from the interview guide and inductive codes from an initial read of the transcripts [50]. We conducted 2 rounds of reliability coding, adjusted the codebook codes and definitions as needed, then completed coding using 1 coder. Coded exports were then organized in interpretation memos to summarize possible explanations for what was happening, including a summary of findings, key distinctions and counterexamples, and further points for consideration [51].

Integration

We used joint display tables [52] and compared quantitative and qualitative results to see where the findings converged, diverged, or expanded [23].

Results

**Participants**

A total of 586 older adults participated in the evaluation over the study period, of whom 289 (49.3%) completed the 6-month follow-up survey (n=25, 8.7% completed the survey over the
telephone). The primary reason for noncompletion was our inability to contact the older adults by telephone or email; only 10 (1.7%) of the 586 older adults declined to participate in the follow-up survey after being contacted.

Table 1 shows remote EBP evaluation participant demographics for those who completed both baseline and 6-month follow-up surveys. The mean age of the participants was 65.4 (SD 12.0) years. The majority were female (241/289, 83.4%) and White (184/289, 63.7%). A little more than one-third (108/289, 37.3%) identified as people of color: American Indian or Alaska Native (2/289, 0.7%), Asian (12/289, 4.2%), Black or African American (77/289, 26.6%), and Latinx (17/289, 5.9%). One in 7 (41/289, 14.2%) lived in a rural area, and 1 in 3 (99/289, 34.3%) found it “somewhat hard” or “very hard” to pay for basics such as food and housing. Of the 289 participants, 113 (39.1%) lived alone, and 58 (20.1%) were providing caregiving. The participants reported a mean of 2.5 (SD 1.7) chronic conditions, with the most common being hypertension (145/289, 50.2%), arthritis (131/289, 45.3%), diabetes (130/289, 45%), and mental health conditions (80/289, 27.7%). The prevalence of all chronic conditions is provided in Table 1, including more rare but still impactful conditions such as Crohn disease and ulcerative colitis.
Table 1. Demographics of remote evidence-based program evaluation participants.

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Total (n=289)</th>
<th>CDSMP(^a) (n=69)</th>
<th>CPSMP(^b) (n=47)</th>
<th>DSMP(^c) (n=118)</th>
<th>EF(^d) (n=12)</th>
<th>WWE(^e) (n=40)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (y), mean (SD)</strong></td>
<td>65.4 (12.0)</td>
<td>60.2 (13.5)</td>
<td>67.5 (12.4)</td>
<td>64.9 (11.1)</td>
<td>72.2 (6.13)</td>
<td>71.2 (8.5)</td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>241 (83.4)</td>
<td>58 (84.1)</td>
<td>39 (83)</td>
<td>93 (78.8)</td>
<td>12 (100)</td>
<td>36 (90)</td>
</tr>
<tr>
<td>Male</td>
<td>48 (16.6)</td>
<td>11 (15.9)</td>
<td>8 (17)</td>
<td>25 (21.2)</td>
<td>0 (0)</td>
<td>4 (10)</td>
</tr>
<tr>
<td>Nonbinary</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>2 (0.7)</td>
<td>1 (1.4)</td>
<td>0 (0)</td>
<td>1 (0.8)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Asian</td>
<td>12 (4.2)</td>
<td>2 (2.9)</td>
<td>2 (4.3)</td>
<td>6 (5.1)</td>
<td>2 (16.7)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>77 (26.6)</td>
<td>15 (21.7)</td>
<td>23 (48.9)</td>
<td>25 (21.2)</td>
<td>1 (8.3)</td>
<td>13 (32.5)</td>
</tr>
<tr>
<td>Native Hawaiian or Pacific Islander</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>White</td>
<td>184 (63.7)</td>
<td>47 (68.1)</td>
<td>19 (40.4)</td>
<td>81 (68.6)</td>
<td>8 (66.7)</td>
<td>26 (65)</td>
</tr>
<tr>
<td><strong>Ethnicity (Latinx), n (%)</strong></td>
<td>17 (5.9)</td>
<td>9 (13)</td>
<td>2 (4.3)</td>
<td>5 (4.2)</td>
<td>1 (8.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Living in a rural area, n (%)</strong></td>
<td>41 (14.2)</td>
<td>12 (17.4)</td>
<td>7 (14.9)</td>
<td>12 (10.2)</td>
<td>3 (25)</td>
<td>6 (15)</td>
</tr>
<tr>
<td>Somewhat hard or very hard to pay for basics, n (%)</td>
<td>99 (34.3)</td>
<td>28 (40.6)</td>
<td>12 (25.6)</td>
<td>45 (38.1)</td>
<td>2 (16.7)</td>
<td>12 (30)</td>
</tr>
<tr>
<td>Living alone, n (%)</td>
<td>113 (39.1)</td>
<td>25 (36.2)</td>
<td>24 (51.1)</td>
<td>40 (33.9)</td>
<td>5 (41.7)</td>
<td>18 (45)</td>
</tr>
<tr>
<td>Caregiver, n (%)</td>
<td>58 (20.1)</td>
<td>14 (20.3)</td>
<td>8 (17)</td>
<td>26 (52)</td>
<td>2 (16.7)</td>
<td>7 (17.5)</td>
</tr>
<tr>
<td>Chronic conditions, mean (SD)</td>
<td>2.5 (1.7)</td>
<td>2.5 (1.5)</td>
<td>2.6 (2.0)</td>
<td>2.7 (1.7)</td>
<td>1.7 (2.0)</td>
<td>1.8 (1.3)</td>
</tr>
<tr>
<td><strong>Chronic conditions(^f), n (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asthma, emphysema, chronic obstructive pulmonary disease, or chronic bronchitis</td>
<td>46 (15.9)</td>
<td>12 (17.4)</td>
<td>11 (23.4)</td>
<td>18 (15.3)</td>
<td>2 (16.7)</td>
<td>3 (7.5)</td>
</tr>
<tr>
<td>Arthritis (rheumatoid arthritis)</td>
<td>18 (6.2)</td>
<td>3 (4.3)</td>
<td>2 (4.3)</td>
<td>8 (6.8)</td>
<td>1 (8.3)</td>
<td>4 (10)</td>
</tr>
<tr>
<td>Arthritis (osteoarthritis)</td>
<td>72 (24.9)</td>
<td>18 (26.1)</td>
<td>6 (12.8)</td>
<td>28 (23.7)</td>
<td>3 (25)</td>
<td>17 (42.5)</td>
</tr>
<tr>
<td>Arthritis (other diagnosis)</td>
<td>41 (14.2)</td>
<td>11 (15.9)</td>
<td>16 (34)</td>
<td>8 (6.8)</td>
<td>0 (0)</td>
<td>6 (15)</td>
</tr>
<tr>
<td>Cancer</td>
<td>15 (5.2)</td>
<td>1 (1.4)</td>
<td>3 (6.4)</td>
<td>9 (7.6)</td>
<td>2 (16.7)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>130 (45)</td>
<td>24 (34.8)</td>
<td>13 (27.7)</td>
<td>84 (71.2)</td>
<td>3 (25)</td>
<td>6 (15)</td>
</tr>
<tr>
<td>Heart trouble (eg, angina, congestive heart failure, and coronary artery disease)</td>
<td>49 (17)</td>
<td>12 (17.4)</td>
<td>7 (14.9)</td>
<td>25 (21.2)</td>
<td>2 (16.7)</td>
<td>3 (7.5)</td>
</tr>
<tr>
<td>Hypertension or high blood pressure</td>
<td>145 (50.2)</td>
<td>29 (42)</td>
<td>27 (57.4)</td>
<td>65 (55.1)</td>
<td>5 (41.7)</td>
<td>19 (47.5)</td>
</tr>
<tr>
<td>Irritable bowel syndrome</td>
<td>18 (6.2)</td>
<td>7 (10.1)</td>
<td>2 (5.9)</td>
<td>7 (5.9)</td>
<td>0 (0)</td>
<td>2 (5)</td>
</tr>
<tr>
<td>Kidney problems</td>
<td>22 (7.6)</td>
<td>6 (8.7)</td>
<td>5 (10.6)</td>
<td>9 (7.6)</td>
<td>1 (8.3)</td>
<td>1 (2.5)</td>
</tr>
<tr>
<td>Liver problems (eg, cirrhosis)</td>
<td>3 (1)</td>
<td>1 (1.4)</td>
<td>0 (0)</td>
<td>2 (1.7)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Mental health conditions (eg, depression, anxiety, posttraumatic stress disorder, and bipolar disorder)</td>
<td>80 (27.7)</td>
<td>28 (40.6)</td>
<td>15 (31.9)</td>
<td>31 (26.3)</td>
<td>0 (0)</td>
<td>6 (15)</td>
</tr>
<tr>
<td>Other digestive problems (besides irritable bowel syndrome, ulcerative colitis, and Crohn disease)</td>
<td>34 (11.8)</td>
<td>10 (14.5)</td>
<td>3 (6.4)</td>
<td>15 (12.7)</td>
<td>1 (8.3)</td>
<td>5 (12.5)</td>
</tr>
<tr>
<td>Stroke and other cerebrovascular disease</td>
<td>13 (4.5)</td>
<td>13 (4.3)</td>
<td>4 (8.5)</td>
<td>6 (5.1)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

\(^{a}\)CDSMP: Chronic Disease Self-Management Program.

\(^{b}\)CPSMP: Chronic Pain Self-Management Program.

\(^{c}\)DSMP: Diabetes Self-Management Program.

\(^{d}\)EF: EnhanceFitness.

\(^{e}\)WWE: Walk With Ease.

\(^{f}\)Less than 1% of the participants reported these chronic conditions: HIV or AIDS, Crohn disease, and ulcerative colitis.
Participants who completed the 6-month follow-up survey had similar demographic characteristics and baseline health status as those who completed only the baseline survey, with a few exceptions. A larger proportion of the CDSMP and CPSMP survey completers identified as Black compared to non–survey completers (15/69, 22% vs 13/98, 13%, and 23/47, 49% vs 31/75, 41%, respectively). WWE survey completers were less likely to be living alone (18/40, 45% vs 24/41, 59%). EF survey completers were more likely to be caregivers than not (2/12, 16% vs 0/10, 0%). Finally, across all programs except EF, survey completers were less likely to be living in a rural area than non–survey completers (37/274, 13.5% vs 109/284, 38.4%).

A total of 198 remote EBP providers (n=123, 62.1% leaders; n=75, 37.9% managers) from 107 EBP organizations in 33 states participated in the evaluation. The majority of the leaders identified as female (113/120, 94.2%). Furthermore, 4.3% (5/117) identified as Asian, 12.8% (15/117) as Black or African American, 11.7% (14/120) as Latinx, and 0.9% (1/117) as biracial. One-quarter of the leaders (31/113, 27.4%) lived in rural settings, one-third (36/120, 30%) were caregivers, and one-third (37/123, 30.1%) had ≥2 chronic conditions. One-fourth (30/121, 24.7%) identified as certified health professionals and 43.3% (52/120) as community health workers, promotoras, or other lay health providers. The leaders had a range of experience in EBP delivery: a little more than half (65/123, 52.8%) had delivered both in-person and remote programs before the survey, while 36.6% (45/123) were conducting remote EBPs for the first time. In addition to completing the survey, 26 EBP administrators, managers, and leaders took part in qualitative interviews. Most of the interview participants (22/26, 85%) identified as female and worked at community or government organizations.

**Outcomes (Quantitative)**

**Overview**

Outcomes are reported by specific program and across the 5 programs included in our evaluation (Tables 2 and 3; Figure 2; Multimedia Appendix 2).
Table 2. Participant health outcomes at enrollment and 6-month follow-up by remote evidence-based program.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Total (n=289), mean (SD)</th>
<th>CDSMP&lt;sup&gt;a&lt;/sup&gt; (n=69), mean (SD)</th>
<th>CPSMP&lt;sup&gt;b&lt;/sup&gt; (n=47), mean (SD)</th>
<th>DSMP&lt;sup&gt;c&lt;/sup&gt; (n=118), mean (SD)</th>
<th>EF&lt;sup&gt;d&lt;/sup&gt; (n=12), mean (SD)</th>
<th>WWE&lt;sup&gt;e&lt;/sup&gt; (n=40), mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health (range 1-5) ↓&lt;sup&gt;f&lt;/sup&gt;</td>
<td>3.23 (0.9)</td>
<td>2.80&lt;sup&gt;b&lt;/sup&gt; (0.9)</td>
<td>3.36 (0.9)</td>
<td>2.91&lt;sup&gt;b&lt;/sup&gt; (1)</td>
<td>3.21 (1.0)</td>
<td>2.71&lt;sup&gt;b&lt;/sup&gt; (0.8)</td>
</tr>
<tr>
<td>Fatigue (range 1-10) ↓</td>
<td>5.49 (2.3)</td>
<td>4.98&lt;sup&gt;b&lt;/sup&gt; (2.3)</td>
<td>6.05 (2.1)</td>
<td>5.61&lt;sup&gt;b&lt;/sup&gt; (2.4)</td>
<td>5.92 (2.3)</td>
<td>4.88&lt;sup&gt;b&lt;/sup&gt; (2.1)</td>
</tr>
<tr>
<td>Pain (range 1-10) ↓</td>
<td>4.86 (2.6)</td>
<td>4.44 (2.7)</td>
<td>5.02 (2.4)</td>
<td>4.78 (2.6)</td>
<td>6.41 (2.3)</td>
<td>5.67&lt;sup&gt;b&lt;/sup&gt; (2.3)</td>
</tr>
<tr>
<td>Sleep quality (range 1-10) ↓</td>
<td>4.95 (2.5)</td>
<td>4.49&lt;sup&gt;b&lt;/sup&gt; (2.6)</td>
<td>5.32 (2.4)</td>
<td>5.07&lt;sup&gt;b&lt;/sup&gt; (2.8)</td>
<td>5.64 (2.6)</td>
<td>5.19 (2.5)</td>
</tr>
<tr>
<td>Loneliness (range 3-9) ↓</td>
<td>4.67 (1.8)</td>
<td>4.45&lt;sup&gt;b&lt;/sup&gt; (1.8)</td>
<td>5.04 (1.9)</td>
<td>4.74&lt;sup&gt;o&lt;/sup&gt; (2.0)</td>
<td>4.57&lt;sup&gt;o&lt;/sup&gt; (1.7)</td>
<td>4.67&lt;sup&gt;o&lt;/sup&gt; (2.0)</td>
</tr>
<tr>
<td>Social isolation (range 5-25) ↑</td>
<td>15.7 (3.7)</td>
<td>16.2 (3.8)</td>
<td>15.2 (3.4)</td>
<td>16.3 (3.5)</td>
<td>15.9 (4.1)</td>
<td>16.7 (4.5)</td>
</tr>
<tr>
<td>Physical activity (day; range 0-77) ↑</td>
<td>2.3 (2.3)</td>
<td>2.6 (2.3)</td>
<td>2.0 (2.1)</td>
<td>2.2 (2.2)</td>
<td>2.5 (2.7)</td>
<td>2.6 (2.5)</td>
</tr>
<tr>
<td>Physical activity (min; range 0-679) ↑</td>
<td>82.5 (113)</td>
<td>98.5&lt;sup&gt;b&lt;/sup&gt; (124)</td>
<td>54.7 (74)</td>
<td>74.7&lt;sup&gt;b&lt;/sup&gt; (106)</td>
<td>86.4 (117)</td>
<td>99.4 (149)</td>
</tr>
<tr>
<td>Depression (range 0-24) ↓</td>
<td>6.20 (5.3)</td>
<td>5.05&lt;sup&gt;o&lt;/sup&gt; (5.2)</td>
<td>7.47 (5.8)</td>
<td>6.07 (5.2)</td>
<td>7.63 (5.6)</td>
<td>6.50 (5.7)</td>
</tr>
<tr>
<td>Anxiety (range 0-6) ↓</td>
<td>1.48 (1.6)</td>
<td>1.30 (1.6)</td>
<td>1.95 (1.7)</td>
<td>1.65 (1.6)</td>
<td>1.67 (1.7)</td>
<td>1.54 (1.6)</td>
</tr>
<tr>
<td>Technology anxiety (range 1-10) ↑</td>
<td>2.74 (2.4)</td>
<td>2.53&lt;sup&gt;b&lt;/sup&gt; (2.4)</td>
<td>3.24 (2.7)</td>
<td>2.81 (2.6)</td>
<td>3.30 (3.1)</td>
<td>2.87 (2.5)</td>
</tr>
<tr>
<td>Technology usability (range 1-10) ↑</td>
<td>7.94 (2.6)</td>
<td>8.22 (2.6)</td>
<td>7.66 (2.5)</td>
<td>8.48 (2.4)</td>
<td>7.15 (3.3)</td>
<td>7.82 (2.9)</td>
</tr>
<tr>
<td>Self-efficacy (range 1-10) ↑</td>
<td>N/A&lt;sup&gt;g&lt;/sup&gt;</td>
<td>N/A</td>
<td>6.16 (2.3)</td>
<td>6.89&lt;sup&gt;b&lt;/sup&gt; (2.2)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Pain interference (range 6-30) ↓</td>
<td>N/A</td>
<td>N/A</td>
<td>16.6 (7.3)</td>
<td>16.7 (7.5)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Diabetes (hypoglycemia; range 1-7) ↓</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<sup>a</sup>CDSMP: Chronic Disease Self-Management Program.  
<sup>b</sup>CPSMP: Chronic Pain Self-Management Program.  
<sup>c</sup>DSMP: Diabetes Self-Management Program.  
<sup>d</sup>EF: EnhanceFitness.  
<sup>e</sup>WWE: Walk With Ease.  
<sup>f</sup>Pre: health outcomes at program enrollment; follow-up: health outcomes at 6-month follow-up from program enrollment.  
<sup>g</sup>Lower scores indicate better health.  
<sup>h</sup>P<.05 (paired t-tests, except for EF, which used the Fisher exact test).  
<sup>i</sup>Higher scores indicate better health.  
<sup>j</sup>N/A: not applicable.  

https://aging.jmir.org/2024/1/e52069
<table>
<thead>
<tr>
<th>Outcome</th>
<th>Range direction</th>
<th>Effect size</th>
<th>Total (n=289)</th>
<th>CDSMP(^a) (n=69)</th>
<th>CPSMP(^b) (n=47)</th>
<th>DSMP(^c) (n=118)</th>
<th>EF(^d) (n=12)</th>
<th>WWE(^e) (n=40)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Health</td>
<td>1-5(^f)</td>
<td>−0.37(^g)</td>
<td>−0.37(^g)</td>
<td>−0.23(^h)</td>
<td>−0.44(^g)</td>
<td>−0.58(^i)</td>
<td>−0.35(^i)</td>
<td></td>
</tr>
<tr>
<td>Fatigue</td>
<td>1-10(^j)</td>
<td>−0.23(^b)</td>
<td>−0.20(^b)</td>
<td>−0.42(^g)</td>
<td>−0.13(^h)</td>
<td>−0.25(^b)</td>
<td>−0.33(^b)</td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>1-10(^k)</td>
<td>−0.13(^b)</td>
<td>−0.11(^b)</td>
<td>−0.49(^g)</td>
<td>0.05(^j)</td>
<td>−0.17(^b)</td>
<td>−0.21(^b)</td>
<td></td>
</tr>
<tr>
<td>Sleep quality</td>
<td>1-10(^l)</td>
<td>−0.18(^b)</td>
<td>−0.19(^b)</td>
<td>−0.08(^i)</td>
<td>−0.21(^h)</td>
<td>−0.66(^i)</td>
<td>−0.22(^b)</td>
<td></td>
</tr>
<tr>
<td>Loneliness</td>
<td>3-9(^m)</td>
<td>−0.13(^b)</td>
<td>−0.24(^b)</td>
<td>0.06(^j)</td>
<td>−0.13(^h)</td>
<td>0.58(^j)</td>
<td>−0.30(^e)</td>
<td></td>
</tr>
<tr>
<td>Social isolation</td>
<td>5-25(^n)</td>
<td>0.08(^i)</td>
<td>0.03(^i)</td>
<td>0.07(^j)</td>
<td>0.00(^j)</td>
<td>0.23(^h)</td>
<td>0.23(^h)</td>
<td></td>
</tr>
<tr>
<td>Physical activity (days)</td>
<td>0-7(^o)</td>
<td>0.13(^h)</td>
<td>0.17(^h)</td>
<td>0.07(^j)</td>
<td>0.10(^j)</td>
<td>0.35(^e)</td>
<td>0.18(^h)</td>
<td></td>
</tr>
<tr>
<td>Physical activity (min)</td>
<td>3-679(^p)</td>
<td>0.15(^b)</td>
<td>0.22(^h)</td>
<td>0.18(^h)</td>
<td>0.08(^j)</td>
<td>0.36(^g)</td>
<td>0.18(^h)</td>
<td></td>
</tr>
<tr>
<td>Depression</td>
<td>0-24(^q)</td>
<td>−0.23(^b)</td>
<td>−0.24(^b)</td>
<td>−0.27(^h)</td>
<td>−0.23(^h)</td>
<td>−0.34(^e)</td>
<td>0.06(^j)</td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>0-6(^r)</td>
<td>−0.08(^i)</td>
<td>−0.12(^b)</td>
<td>−0.36(^g)</td>
<td>0.03(^j)</td>
<td>0.00(^j)</td>
<td>−0.06(^i)</td>
<td></td>
</tr>
<tr>
<td>Technology anxiety</td>
<td>1-10(^s)</td>
<td>−0.13(^b)</td>
<td>−0.17(^b)</td>
<td>−0.23(^h)</td>
<td>−0.08(^j)</td>
<td>−0.03(^j)</td>
<td>−0.24(^b)</td>
<td></td>
</tr>
<tr>
<td>Technology usability</td>
<td>1-10(^t)</td>
<td>−0.05(^i)</td>
<td>−0.06(^i)</td>
<td>0.23(^h)</td>
<td>−0.13(^h)</td>
<td>0.59(^j)</td>
<td>−0.07(^j)</td>
<td></td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>1-10(^u)</td>
<td>N/A(^l)</td>
<td>0.38(^g)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Pain interference</td>
<td>6-30(^v)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>0.37(^g)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>Diabetes (hypoglycemia)</td>
<td>1-7(^w)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>−0.06(^j)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) CDSMP: Chronic Disease Self-Management Program.  
\(^b\) CPSMP: Chronic Pain Self-Management Program.  
\(^c\) DSMP: Diabetes Self-Management Program.  
\(^d\) EF: EnhanceFitness.  
\(^e\) WWE: Walk With Ease.  
\(^f\) Lower scores indicate better health.  
\(^g\) Cohen \(d\) effect sizes 0.3 to 0.5 = moderate.  
\(^h\) Cohen \(d\) effect sizes 0.1 to 0.3 = small.  
\(^i\) Cohen \(d\) effect sizes > 0.5 = large.  
\(^j\) Cohen \(d\) effect sizes < 0.1 = trivial.  
\(^k\) Higher scores indicate better health.  
\(^l\) N/A: not applicable.
**Figure 2.** Change in the number of physical activity (PA) days between enrollment and 6-month follow-up for remote evidence-based program outcome evaluation participants. The participants were asked about the number of days on which they were physically active per week at both program enrollment and 6-month follow-up. The completion rates for this question, by program and overall, are as follows: Walk With Ease (WWE)=90% (36/40), EnhanceFitness (EF)=100% (12/12), Diabetes Self-Management Program (DSMP)=87.3% (103/118), Chronic Pain Self-Management Program (CPSMP)=45% (21/47), Chronic Disease Self-Management Program (CDSMP)=94% (65/69), and overall=82% (237/289).

**CDSMP Participants**

When testing for whether these changes over time were not due to chance, CDSMP participants showed statistically significant improvements in health (mean change 0.32, 95% CI 0.117-0.530; *P*=.003; *t* =3.12), energy (fatigue; mean change 0.73, 95% CI 0.061-1.408; *P*=.03; *t* =2.18), sleep quality (mean change 0.89, 95% CI 0.216-1.561; *P*=.01; *t* =2.64), and self-efficacy (mean change −0.53, 95% CI −0.970 to −0.091; *P*=.02; *t* =2.49). Participants with depression at baseline (19/69, 28%; PHQ-8 score <10) also significantly reduced their depression symptom severity from mean 14.7 (SD 4.2) to mean 11.5 (SD 5.1) on the PHQ-8 (*P*=.03). Overall, 58% (40/69) of the participants improved their self-efficacy over time, with a moderate effect size of 0.38. From program enrollment to 6-month follow-up, 40% (26/65) of the participants showed improvement, and 34% (22/65) of the participants maintained the number of days on which they were physically active.

**CPSMP Participants**

CPSMP participants significantly improved their energy (fatigue; mean change 1.11, 95% CI 0.248-1.968; *P*=.01; *t* =2.61) and pain (mean change 0.97, 95% CI 0.186-1.763; *P*=.02; *t* =2.50). Moderate effect sizes were seen for fatigue (−0.42), pain (−0.49), anxiety (−0.36), and pain interference (−0.38). The participants with pain interference at baseline (30/289, 10.4%; PROMIS [Patient-Reported Outcomes Measurement Information System] score >6) also reduced their pain interference from mean 16.9 (SD 6.6) to mean 15.5 (SD 7.3). Only 1 (2.1%) of the 47 participants was taking opioid medications at enrollment. From program enrollment to 6-month follow-up, 29% (6/21) of the participants showed improvement, and 43% (9/21) of the participants maintained the number of days on which they were physically active.

**DSMP Participants**

DSMP participants demonstrated better health (mean change 0.41, 95% CI 0.259-0.552; *P*<.001; *t* =5.48). People with at least 1 hypoglycemia symptom at baseline (70/289, 24.2%) reduced their symptoms from mean 2.4 (SD 1.3) to mean 2.1 (SD 1.7); the effect size was low (−0.06). DSMP participants had a moderate effect size for overall improvement in health (−0.44). Participants with depression at baseline (19/118, 16.1%) also significantly reduced their depression symptom severity from mean 13.9 (SD 3.0) to mean 11.1 (SD 6.5) on the PHQ-8. From program enrollment to 6-month follow-up, 39.8% (41/103) of the participants showed improvement, and 33% (34/103) of the participants maintained the number of days on which they were physically active.

**EF Participants**

EF participants demonstrated better sleep quality (Fisher exact test: *P*=.001), with a large effect size of −0.66. Participants also showed strong improvements in health (effect size: −0.58) and technology usability (effect size: 0.60). From program enrollment to 6-month follow-up, 42% (5/12) of the participants showed improvement, and 33% (4/12) of the participants
maintained the number of days on which they were physically active.

**WWE Participants**

WWE participants demonstrated better health (mean change 0.25, 95% CI 0.046–0.454; \( P=0.02; t_{35}=-2.49 \)), with a moderate effect size of \(-0.35\). Participants also had moderate effect sizes for improvements in fatigue (\(-0.33\)) and loneliness (\(-0.30\)). From program enrollment to 6-month follow-up, 36% (13/36) of the participants showed improvement, and 33% (12/36) of the participants maintained the number of days on which they were physically active.

**Outcomes (Qualitative)**

**Participants**

Overall, participants shared that they liked and enjoyed participating in the remote EBPs and identified several ways in which the program impacted their lives. First, they reported changes in their knowledge, attitudes, and practices in promoting their health and well-being. Participating in remote EBPs helped older adults manage their chronic conditions, become more active, and feel more confident and better equipped to take care of themselves:

*I feel as if this program literally changed the trajectory of my life. Prior to it, I was diagnosed as prediabetic and was put on medication, which made me very sick. My next option was a very expensive diabetes drug. But, through following this program, I learned about insulin resistance and what I could do to combat my descent into diabetes. I was encouraged and coached in inspiring ways. I am now barely considered even prediabetic.* [Female DSMP participant #1454; aged 66 years; living with ulcerative colitis]

*The program helped me understand how exercise can improve my mobility, and it encouraged me to remain active.* [Male WWE participant #1346; aged 63 years; living with arthritis and asthma]

Likewise, the remote EBP participants stated that they learned a lot from the program, including from the materials, leaders, and other participants. This was true across programs: WWE participants learned about new walking spots in their communities from other walkers in their cohort, while self-management program participants learned new ways and tips for managing their condition (expanded their sense of their own options) and broadened their understanding of what life was like at different severity levels of their condition. The group-based formats of the remote EBPs helped provide accountability and motivation and also provided a variety of perspectives and ideas. For some participants, the remote program helped them accept the reality and seriousness of their condition and the changes needed to manage it. In addition, participants believed that others could benefit from the program as well.

The participants also reported social benefits from being part of the programs. Many participants felt less alone, gained a sense of comfort from talking with others struggling with the same conditions, and made friends over the course of the program. A sense of camaraderie was reported often:

*I liked the interaction with other people, it’s helpful to find out how others are going through. To know that there’s other people out there with a lot of pain and they’re struggling with it, made me feel not alone. Some are worse and some are better. It feels isolating a lot with pain, so that was really nice to experience, seeing others.* [Female CPSMP participant #1344; aged 74 years; living alone and managing multiple chronic conditions]

Furthermore, remote EBP participants from various programs who were grappling with new or existing chronic conditions, changes in mobility and function due to aging, and the challenges of physical distancing during the pandemic, emphasized the value of learning together and feeling less alone. Even participants in the remote WWE program (self-directed tool kit enhanced with a leader virtually supporting a cohort of participants) found social benefits:

*I did appreciate the opportunity to meet with the group assigned to me and get the encouragement to get out and walking.* [Female WWE participant #1249; aged 65 years; living with arthritis and a mental health condition]

In addition, older adults shared how remote EBPs improved their access to the programs. The remote format made accessing these programs during the pandemic both safe and very convenient, in particular for people living with chronic pain or disabilities. Some participants would not have been able to participate if the programs had only been offered in person. That said, a few participants did not like the remote format and found it more difficult to access. For these participants, the

Across programs, remote EBP participants showed statistically significant improvements in their health (mean change 0.33, 95% CI 0.235–0.422; \( P<0.01; t_{260}=6.92 \)), energy (mean change 0.56, 95% CI 0.264–0.853; \( P<0.01; t_{264}=3.73 \)), sleep quality (mean change 0.53, 95% CI 0.245–0.812; \( P<0.01; t_{265}=3.67 \)) loneliness (mean change 0.25, 95% CI 0.057–0.437; \( P=0.01; t_{242}=2.55 \)), depressive symptoms (mean change 0.60, 95% CI 0.111–1.091; \( P=0.02; t_{212}=2.42 \)), and technology anxiety (mean change 0.34, 95% CI 0.010–0.665; \( P=0.04; t_{231}=2.03 \)). For people living with clinically significant depressive symptoms (PHQ-8 score ≥10; 24/114, 21.1% of the sample), the overall mean change in the PHQ-8 score from enrollment to 6-month follow-up was 3.025 (95% CI 1.379–4.671; \( P<0.01; t_{260}=3.717 \)). People with clinically significant depression at baseline (52/289, 18%) also significantly reduced their depression symptom severity from mean 14.7 (SD 4.2) to mean 11.5 (SD 5.1).

The percentage change is reported in Multimedia Appendix 2. Across programs, 26.7% (66/247) to 49.8% (136/273) of the participants improved health outcomes over time, and 17.4% (41/235) to 53.8% (133/247) of the participants maintained their improvements. The percentage change is reported in Multimedia Appendix 2.
downsides of not being able to meet and connect in person or the challenges with participating via telephone, mail, or videoconferencing outweighed the benefits of participating in class. Examples include issues with the technology itself (eg, poor internet connections that made the videoconferencing software freeze up) and discomfort with using technology (eg, unfamiliarity with navigating Zoom functions or unease using a mobile phone where they cannot see other participants). It should be noted that technology encompasses using a telephone (landline, smartphone, or other mobile phone) in addition to laptop computers, tablet devices, and PCs; dial-up or broadband internet; and videoconferencing platforms such as Zoom and Webex.

Finally, many participants shared that they wanted to take part in the remote EBP again, and some participants shared that they desired more follow-ups after the program ended. For these participants, there was a sense of having missed or forgotten some of what had been taught and wanting to refresh their knowledge. Others felt that they needed the motivation of continued check-ins to keep using what they had learned. In addition, some were not sure how to get their questions answered after the program ended, how to sign up for other programs offered by the organization, or whether they were allowed to take part in the program again. This points to opportunities for future supports and services after remote EBP engagement, such as monthly check-ins via telephone, videoconferencing, or social media to “keep the feelings of motivation and community after the program ends...[to] meet or discuss what folks are doing and what works and encourage each other to keep going” (female DSMP participant #1190; aged 61 years; living with multiple chronic conditions [arthritis, hypertension, diabetes, heart troubles, and a mental health condition]). Post-EBP supports and services could also provide a way to reinforce and deepen knowledge and skills that are learned and practiced during a relatively brief program and offer ways for family, friends, and caregivers to support the maintenance of program gains as well as widen program benefits to other people in the participants’ communities.

**Providers**

Table 4 summarizes provider’s perspectives on the impact of remote EBPs. Some of these impacts were expected; for example, 3 in 4 providers (149/198, 75.2%) reported improved health outcomes for older remote EBP participants, and reducing social isolation and loneliness emerged as the most common benefit for both participants and EBP providers (leaders and managers). In addition, half of the providers (102/198, 51.5%) noted that connections to other supports and services were a participant benefit; while this typically occurs during in-person EBP delivery, remote EBP delivery allowed for sharing timely and ever-changing information about testing for COVID-19 infection and recommended safety protocols, as well as referrals to services that may have paused or been shifted due to pandemic-related closures or physical distancing requirements.
Table 4. Perceived benefits of remote evidence-based program (EBP) delivery for participants (from providers’ perspectives) and for providers (n=198).

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Survey data, n (%)</th>
<th>Interview data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Leaders (n=123)</td>
<td>Managers (n=75)</td>
</tr>
<tr>
<td>Benefits for participants (from providers’ perspectives)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved health outcomes</td>
<td>93 (75.6)</td>
<td>56 (74.7)</td>
</tr>
<tr>
<td></td>
<td>Kept participants safe while also allowing them to access the benefits of these programs.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The program was still effective, and the benefit it had on participants could be seen.</td>
<td></td>
</tr>
<tr>
<td>Reduced social isolation and loneliness</td>
<td>116 (94.3)</td>
<td>66 (88)</td>
</tr>
<tr>
<td></td>
<td>Enhanced socialization and helped with isolation, which is really needed right now. The bonds formed within the groups are really important.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clients appreciated getting checked on; many were feeling isolated during the COVID-19 pandemic.</td>
<td></td>
</tr>
<tr>
<td>Improved access to technology</td>
<td>52 (42.3)</td>
<td>30 (40)</td>
</tr>
<tr>
<td></td>
<td>By participating in the remote EBP, participants gained access to new or loaner technology.</td>
<td></td>
</tr>
<tr>
<td>Improved comfort with using technology</td>
<td>91 (74)</td>
<td>52 (69.3)</td>
</tr>
<tr>
<td></td>
<td>Increased technology literacy and comfort of participants, which encourages them to explore other web-based resources.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clients felt accomplished to have completed a class that required new technology, such as video-conferencing platforms, without help.</td>
<td></td>
</tr>
<tr>
<td>Enhanced access to other supports and services</td>
<td>67 (54.5)</td>
<td>35 (46.6)</td>
</tr>
<tr>
<td></td>
<td>Able to educate older adults in their program about COVID-19 vaccines.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Remote program improved cross-referrals, which is good for holistically addressing health.</td>
<td></td>
</tr>
<tr>
<td>Improved access to EBPs</td>
<td>N/Aa</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Participants can repeat the program because it is easier to access.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No concerns about driving in bad weather.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Some participants liked the virtual class and want remote options in the future.</td>
<td></td>
</tr>
<tr>
<td>Benefits for providers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Improved health outcomes</td>
<td>58 (47.2)</td>
<td>27 (36)</td>
</tr>
<tr>
<td></td>
<td>Able to keep working and connecting with colleagues and participants.</td>
<td></td>
</tr>
<tr>
<td>Reduced social isolation and loneliness</td>
<td>61 (49.6)</td>
<td>39 (52)</td>
</tr>
<tr>
<td></td>
<td>Gained access to technology via work or family, friends, and neighbors.</td>
<td></td>
</tr>
<tr>
<td>Improved access to technology</td>
<td>34 (27.6)</td>
<td>22 (29.3)</td>
</tr>
<tr>
<td></td>
<td>For leaders without access, some stopped delivering the program.</td>
<td></td>
</tr>
</tbody>
</table>

aN/A: not applicable (either the benefit was not one of the multiple-choice answer choices in the provider survey or did not emerge during the interview data analysis).

Some of the impacts were positive but unintended; for example, providers reported improved comfort using technology as a benefit for remote EBP participants. Increasing the usability of technology and reducing anxiety about technology are not original outcomes that in-person EBPs strived to impact, but in remote EBP delivery, opportunities arose for some participants to become more comfortable using technology (telephone, videoconferencing, or tool kits) for engaging in, and receiving, other supports and services. Comfort using technology was a more prevalent impact than access to technology, which aligns with other findings that access was a challenge even when organizations provided software or hardware [53]. Other unexpected positive impacts from the providers’ perspectives include that delivering EBPs remotely allowed them to continue providing the program to older adults throughout the COVID-19 pandemic, reach participants they had not reached before, connect participants to each other, learn how to use technology, and be able to continue working or volunteering during the pandemic.

We also gathered data via surveys and interviews on the unintended negative consequences of delivering EBPs remotely, an important aspect when evaluating the public health impact of these programs with an equity lens. For some leaders, teaching remotely was too difficult or disconnecting due to not being able to see people’s nonverbal cues and having to work harder to teach technology, engage people, or address emotional issues such that they felt that “something was lost.” Strategies for mitigating this reduced impact included reducing class sizes or duration (for both telephone and videoconferencing sessions) and requiring people attending videoconferencing sessions to
have their camera on and having 2 leaders taking part so that one could focus on engagement while one managed technology.

**Discussion**

**Principal Findings**

Our evaluation found that remote EBPs showed improvements on various outcomes from program enrollment to 6-month follow-up, including their self-rated health, energy, sleep quality, loneliness, depressive symptoms, and technology anxiety, within and across programs. Some program participants also reduced their anxiety, pain, pain interference, physical activity and self-efficacy, and enhanced their technology usability. The quantitative findings yielded mainly small effect sizes. This may be due to the heterogeneity of the participants; in particular, participants who did not enter the program as lonely or inactive would have no room to improve over time. The qualitative findings suggest that remote EBPs participants improved knowledge, attitudes, and skills on how to live healthier lives; reduced social isolation and loneliness; and gained better access to programs. In addition, providers shared that they too benefited from delivering programs remotely by staying connected, having access to technology, and improving their own health and well-being.

In some cases, the findings support previous research on the effectiveness of the remote EBPs that were part of our evaluation; for example, a study of 213 videoconferencing CDSMP participants in rural and remote Ontario, Canada, found similar improvements in self-rated health, energy, and psychological well-being (a measure related to depressive symptoms) 4 months after their last class [54]. Another study of 97 telephone plus tool kit CPSMP participants in Cleveland, Ohio, United States, also reported better pain outcomes immediately after program completion [17]. Furthermore, a study of self-directed WWE participants (n=270) in rural and urban North Carolina, United States, who were living with arthritis found that participants also reduced fatigue at follow-up 1 year after program enrollment [21]. Our findings are also comparable to those of similar health promotion programs, such as increased physical activity and reduced depression for a remote DSMP [55].

In other cases, our findings contrast with those of previous research; for instance, 1 study of remote-delivered EF [20] found that the participants (n=15) decreased their knee pain as measured by the Knee Injury and Osteoarthritis Outcome Score; however, all study participants had symptomatic knee osteoarthritis, and the postprogram outcome was measured directly after active intervention at 12 weeks. It may be that our evaluation participants did not significantly improve pain outcomes due to their less severe baseline pain. In addition, the study of CPSMP participants mentioned previously [17] found statistically significant improvements in sleep and depression; our evaluation participants too improved their sleep quality and reduced their depressive symptom severity, but this change was not statistically significant. Our differing findings may be due to our smaller sample size (47 vs 97), which lowered our power to detect significant change, or a longer follow-up period (6 months vs 6 weeks), during which improvements may have been attenuated.

Furthermore, it should be noted that our outcome evaluation was not designed to compare in-person EBP delivery with remote EBP delivery. We used a common set of outcome measures of interest to older adults, EBP delivery organizations, and policy makers; as such, some of these measures look at similar constructs as the in-person EBP effectiveness studies but use different instruments, and some of these measures that are newly being evaluated (eg, loneliness and depression) were not evaluated in the in-person EBP research studies. In addition, we used a different time period compared to studies that evaluated the effectiveness of in-person EBP delivery: we looked at changes between program enrollment and 6-month follow-up, regardless of program duration, whereas prior effectiveness research looked at pre-post change over time based on program duration [21,56-59]. Our sample sizes were also smaller than those of previous in-person EBP research studies; therefore, the lack of statistical significance may be due to a lack of power to detect change. That said, it may be helpful to contextualize our remote EBP outcome evaluation findings with those from in-person EBP outcome evaluations.

Remote CDSMP evaluation participants reported improvements in self-efficacy, health, energy (fatigue), sleep quality, physical activity, and depression, all outcomes that were also reported in previous studies of in-person CDSMP [37,60-62]. Furthermore, remote CDSMP evaluation participants did not show the significant improvements in pain outcomes that were seen in research with in-person program participants. Remote CPSMP evaluation participants reported outcome improvements similar to those of in-person CPSMP participants regarding pain and pain interference [56,63], while our evaluation added evidence for remote CPSMP participant improvements in energy (fatigue), anxiety, and physical activity; these 3 outcomes were not assessed in in-person CPSMP studies. However, remote CPSMP participants did not report significant improvements in self-efficacy, as was reported in previous in-person CPSMP studies. Looking at the DSMP, both remote and in-person delivery participants showed reduced hypoglycemia symptoms, health, and depression [57]. Earlier research with in-person program participants found improved self-efficacy as well, and remote program participants showed improvement in terms of their physical activity days. For EF, both remote and in-person program participants reported improvements in self-rated health as well as physical activity [64,65]. This evaluation adds novel findings on improved sleep quality and the usability of technology for remote program participants. In addition, in-person EF participants have reported reduced depression in other studies [64]. Looking at WWE, both in-person and remote program participants have demonstrated improvements in health, fatigue, and physical activity [21]. Remote WWE participants reduced their loneliness, which was not assessed in previous studies of in-person WWE [21]. Previous research of in-person WWE also found participant improvement in pain and self-efficacy [21].

Although EBPs shifting to remote delivery was necessitated by the pandemic-related lockdown and other restrictions [24], the findings suggest that remote delivery can improve access to
quality health promotion programs outside the COVID-19 pandemic context; for example, a caregiver for a person living with dementia can join a physical activity program from home to support their own health while not having to find respite care for their loved one, an older adult living with chronic pain can join a class even if they are experiencing elevated pain levels that would make it hard to leave even their bedroom, and a program leader can deliver the program across a region in Spanish to participants who may not have access to language-specific health promotion. This is similar to other studies of remote EBPs (eg, the value for people living with cancer to be able to participate from a distance despite living with weakened immune systems [66]).

Adapting the format and channels through which an EBP is delivered [67] is considered an appropriate modifiable intervention characteristic to better align with the needs and preferences of participants, as well as organizational and geographic contexts in which a remote option improves program feasibility and sustainability [68]. Implementation science increasingly recognizes the need for adaptations [69] to improve the intervention-context fit (eg, increasing both practical and value fit [70]; increasing EBP feasibility to a given context [71]; and making sure that the EBPs can be delivered for different systems, organizations, providers, and participants than was originally studied [72]), support people living with multiple chronic conditions because comorbidity is the rule rather than the exception [73], and ensure EBP effectiveness by evaluating the intervention in different settings with varying provider and participant attributes [74]. Our complementary process evaluation found that delivering in-person EBPs remotely did not require modifications to core program components [53], suggesting that fidelity to the active ingredients of the program models was maintained. That said, some of the observed lack of improvement in health outcomes by remote EBP participants may be due to failures in implementation (eg, inadequate intervention dose due to technology issues) that negatively impacted program effectiveness [42].

Strengths and Limitations

Our evaluation comes with several strengths. First, we conducted a pragmatic evaluation across the country in partnership with policy makers and delivery organizations so that the findings would have direct implications for improving older adults’ health equity. Specifically, we built upon trusting relationships and took time to foster new relationships to engage people most impacted by implementing remote EBPs; we also gathered qualitative data to put quantitative outcomes in context for future quality improvement [75]. Second, we included multiple brief, validated health outcomes across various EBPs and diverse participants and providers, centering outcomes that are important and meaningful to participants, providers, organizations, and policy makers that were not measured in previous studies [76]. These include social factors such as social isolation and loneliness and mental health factors such as depression and anxiety, which lead to premature mortality for older adults [77,78]. While measuring multiple outcomes across heterogeneous populations and programs made it harder to see large effect sizes, it better reflects how organizations work (delivering multiple programs) and how participants view health more holistically rather than as just 1 primary outcome. It may also be that small effect sizes suggest that these programs provided primary or secondary prevention to delay the onset of more impairing symptoms and conditions that was not picked up from our measures or relatively study time frame [79]. Third, collecting qualitative as well as quantitative data from the perspectives of both participants and providers provides consistent measures to compare across studies as well as stories and unanticipated outcomes to explore more broadly in future research.

Looking at limitations, first, using a single-group design with the lack of a comparison group or randomization means that we cannot attribute change in health outcomes over time to participating in the remote EBP. Second, we recognize that the lack of statistical significance when assessing for whether change was due to chance or program participation may be due to small sample sizes that lack power. This is partly why we collected qualitative data from both providers and participants: we wanted to understand the magnitude of change from their perspectives. Third, our response rate was lower than is typical in controlled research studies, with only half of the remote EBP participants (289/586, 49.3%) providing follow-up data. This was due to pandemic-related logistical and methodological challenges faced by our evaluation team, program providers, and older adult participants; conducting an evaluation during the COVID-19 pandemic has been deemed more challenging than conducting evaluations in conflict areas [80]. This suggests limitations to both internal validity through biases such as selection bias (eg, persons who liked the program or had access to a mobile phone or the internet to complete the evaluation were overrepresented) and external validity, with the COVID-19 pandemic being a unique context, because of which our findings may not be generalizable (eg, were improvements in social connection due to the high level of disconnectedness faced by everyone during the pandemic-related lockdown?). Finally, our convenience sample only included group-based programs and as such may not be applicable to one-on-one EBPs. We also did not have access to data on the entire remote EBP delivery population; therefore, we cannot comment on the representativeness of our outcome evaluation sample. Our process evaluation [24] suggests that persons who have been historically underserved by EBPs (eg, people of color, those with disabilities, and those living in rural areas) can be reached through remote delivery; however, care must be taken to ensure that remote delivery does not widen the health inequities caused by the digital divide.

Conclusions

In closing, the findings from this outcome evaluation suggest that participating in remote EBPs can improve health, social, and technology outcomes of interest for older adults, providers, and policy makers (Textbox 1). Future policy and practice can better support remote EBPs, improving access for all older adults (regardless of income, geography, and ability) and strengthening delivery organizations.
Textbox 1. Lessons learned for evidence-based program (EBP) administrators, developers, and implementers.

**Key takeaways**

- Delivering EBPs remotely (by telephone, videoconferencing, mail, or a combination of these modes) offers a new format for engaging older adults in quality health promotion programming.
- Outcome evaluation findings suggest that remote EBPs are effective at improving health, social, and technology outcomes for older adult participants and providers who deliver the programs, although impacts are not experienced universally across programs or outcomes.
- Interviews and surveys with remote EBP participants and providers suggest that these outcomes can be achieved because remote EBPs are acceptable; increase knowledge, skills, motivation, support, and accountability; connect people to peers and leaders; and support tech access and comfort.
- Include brief, validated pre- and postprogram surveys in your routine program delivery to understand changes in outcomes that matter to your community, organization, and funders. (The surveys may include traditional health outcomes you assessed with in-person EBPs as well as new outcomes, such as social isolation, loneliness, technology anxiety, and technology usability. Our survey measures are cited in this paper and available on request.)
- Gather data on remote EBP participant demographics to understand who is being reached and who is not being reached by remote programs to strategize engagement via other outreach strategies or in-person EBP modes.
- Remote EBP delivery may improve access to health promotion programming for people facing inequitable access to in-person programming; however, caution is needed to ensure that remote delivery does not widen the digital divide; for example, while remote delivery can support rural-dwelling older adults from areas without nearby programs or dependable transportation, the lack of reliable internet in some rural areas may necessitate a telephone or mail remote EBP delivery mode.
- Supports for remote EBPs include orienting participants to using program technology before and during the program and providing ways for program participants to connect after the program ends.

**Acknowledgments**

The authors thank the remote evidence-based program participants, leaders, and staff who came together during this difficult pandemic to continue health promotion from a distance. Study data were collected and managed using REDCap (Research Electronic Data Capture) tools hosted at the Institute of Translational Health Sciences and supported by the National Center for Advancing Translational Sciences of the National Institutes of Health (UL1 TR002319). This evaluation would not have been possible without generous support from the funders. This project was supported by a cooperative agreement (NU38OT000315) funded to the Society for Public Health Education by the National Center for Chronic Disease Prevention and Health Promotion (Division of Population Health, Arthritis Management and Well-Being Program) of the Centers for Disease Control and Prevention; by the Administration for Community Living (RTI contract 37-312-0214780 to Sound Generations); and by the RRF Foundation for Aging (grant 2020183 to Sound Generations). The contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention, Administration for Community Living, RRF Foundation for Aging, or the Department of Health and Human Services.

**Conflicts of Interest**

PD and KL work for organizations that administer remote evidence-based programs. All other authors declare no other conflicts of interest.

**Multimedia Appendix 1**

Participant outcomes survey.

[PDF File (Adobe PDF File), 142 KB - aging_v711e52069_app1.pdf ]

**Multimedia Appendix 2**

The percentages of improvement and maintenance in participant health outcomes at enrollment and 6-month follow-up in remote evidence-based programs for health promotion.

[DOCX File, 45 KB - aging_v711e52069_app2.docx ]

**References**

2. The Longevity Economy® Outlook: how people age 50 and older are fueling economic growth, stimulating jobs, and creating opportunities for all. AARP. 2019. URL: https://www.aarp.org/content/dam/aarp/research/surveys_statistics/econ/2019/longevity-economy-outlook.doi.10.26419-2Ffin.00042.001.pdf [accessed 2022-12-18]


### Abbreviations

ACL: Administration for Community Living  
CDC: Centers for Disease Control and Prevention  
CDSMP: Chronic Disease Self-Management Program  
CPSMP: Chronic Pain Self-Management Program  
DSMP: Diabetes Self-Management Program  
EBP: evidence-based program  
EF: EnhanceFitness  
PHQ: Patient Health Questionnaire  
PROMIS: Patient-Reported Outcomes Measurement Information System  
RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance  
REDCap: Research Electronic Data Capture  
WWE: Walk With Ease
provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Original Paper

Self-Explainable Graph Neural Network for Alzheimer Disease and Related Dementias Risk Prediction: Algorithm Development and Validation Study

Xinyue Hu1,2*, MS; Zenan Sun2*, MS; Yi Nian2, MS; Yichen Wang3, MD; Yifang Dang2, MS; Fang Li1,2, PhD; Jingna Feng1,2, MS; Evan Yu2, MS; Cui Tao1,2, PhD

1Department of Artificial Intelligence and Informatics, Mayo Clinic, Jacksonville, FL, United States
2McWilliams School of Biomedical Informatics, The University of Texas Health Science Center at Houston, Houston, TX, United States
3Division of Hospital Medicine at Perelman School of Medicine, The University of Pennsylvania, Philadelphia, PA, United States

*These authors contributed equally

Corresponding Author:
Cui Tao, PhD
Department of Artificial Intelligence and Informatics
Mayo Clinic
4500 San Pablo Rd S
Jacksonville, FL, 32224
United States
Phone: 1 904 956 3256
Email: tao.cui@mayo.edu

Abstract

Background: Alzheimer disease and related dementias (ADRD) rank as the sixth leading cause of death in the United States, underlining the importance of accurate ADRD risk prediction. While recent advancements in ADRD risk prediction have primarily relied on imaging analysis, not all patients undergo medical imaging before an ADRD diagnosis. Merging machine learning with claims data can reveal additional risk factors and uncover interconnections among diverse medical codes.

Objective: The study aims to use graph neural networks (GNNs) with claim data for ADRD risk prediction. Addressing the lack of human-interpretable reasons behind these predictions, we introduce an innovative, self-explainable method to evaluate relationship importance and its influence on ADRD risk prediction.

Methods: We used a variationally regularized encoder-decoder GNN (variational GNN [VGNN]) integrated with our proposed relation importance method for estimating ADRD likelihood. This self-explainable method can provide a feature-important explanation in the context of ADRD risk prediction, leveraging relational information within a graph. Three scenarios with 1-year, 2-year, and 3-year prediction windows were created to assess the model’s efficiency, respectively. Random forest (RF) and light gradient boost machine (LGBM) were used as baselines. By using this method, we further clarify the key relationships for ADRD risk prediction.

Results: In scenario 1, the VGNN model showed area under the receiver operating characteristic (AUROC) scores of 0.7272 and 0.7480 for the small subset and the matched cohort data set. It outperforms RF and LGBM by 10.6% and 9.1%, respectively, on average. In scenario 2, it achieved AUROC scores of 0.7125 and 0.7281, surpassing the other models by 10.5% and 8.9%, respectively. Similarly, in scenario 3, AUROC scores of 0.7001 and 0.7187 were obtained, exceeding 10.1% and 8.5% than the baseline models, respectively. These results clearly demonstrate the significant superiority of the graph-based approach over the tree-based models (RF and LGBM) in predicting ADRD. Furthermore, the integration of the VGNN model and our relation importance interpretation could provide valuable insight into paired factors that may contribute to or delay ADRD progression.

Conclusions: Using our innovative self-explainable method with claims data enhances ADRD risk prediction and provides insights into the impact of interconnected medical code relationships. This methodology not only enables ADRD risk modeling but also shows potential for other image analysis predictions using claims data.

(JMIR Aging 2024;7:e54748) doi: 10.2196/54748

https://aging.jmir.org/2024/1/e54748

(JMIR Aging 2024 | vol. 7 | e54748 | p.513)
KEYWORDS
Alzheimer disease and related dementias; risk prediction; graph neural network; relation importance; machine learning

Introduction

Background
Alzheimer disease and related dementias (ADRD) currently rank as the sixth leading cause of death in the United States [1]. Currently, 47 million people live with ADRD globally [2]. By the year 2050, the prevalence of dementia is expected to triple worldwide [3]. These alarming statistics emphasize the pressing need for accurately predicting ADRD risk, which holds immense significance for several reasons. First, it enables early detection and diagnosis, which can facilitate timely interventions and treatment plans that have the potential to slow down disease progression, improve patient outcomes, and enhance the quality of life for individuals affected by ADRD. Second, it also plays a crucial role in advancing research and drug development. It provides valuable insights into disease progression, risk factors, and potential therapeutic targets. By identifying individuals at high risk of developing ADRD, researchers can conduct targeted studies and clinical trials and explore preventive measures to mitigate the impact of this debilitating disease. Third, early prediction and intervention may help reduce health care costs associated with ADRD. By identifying individuals at risk and providing appropriate care, the burden on the health care system can be lessened. Nevertheless, predicting ADRD risks is an intricate task due to its nature as a long-term chronic disease with multifaceted underlying causes.

In the context of ADRD risk prediction, the conventional approach predominantly involves using machine learning (ML) models with medical imaging data as primary resources to achieve commendable success [4-6]. However, it is important to acknowledge that not all patients undergo routine clinical imaging tests during their regular visits, rendering medical imaging data less accessible for certain individuals. In contrast, claims data provide a more readily available data source for the ML predictors. Hence, the development of a valuable and easily trainable risk prediction tool necessitates the use of existing claims data as the primary input for prediction. This approach not only enhances the model’s generalizability but also facilitates its adaptation to other diverse data sources.

In recent years, the emergence of graph-structured data has received significant interest within the realm of deep learning [7-11]. Graphs are composed of nodes and relationships, resulting in the representation and analysis of intricate connections and patterns within the data they encapsulate. They also offer a unique combination of topological structure and individual features, which enables a rich source of information [12,13]. To analyze and model the complex relations of interconnected graph data, graph neural networks (GNNs) have emerged as a powerful tool [14]. Unlike traditional ML models that operate on fixed-dimensional inputs, GNNs operate directly on the graph structure, which allows them to learn the representation of individuals, attributes, and relationships. In the biomedical domain, GNNs have been used for tasks such as protein function prediction, drug discovery, disease classification, and personalized medicine [15-20]. Li et al [21] proposed a multi-channel GNN for predicting drug-target interactions that combines a multi-channel graph convolutional network and graph attention network (GAT). This framework uses a topology graph for contextual representation, a feature graph for semantic representation, and a common representation of drug and protein pairs. It has demonstrated remarkable accuracy in identifying drug-target interactions, achieving an impressive area under the receiver operating characteristic (AUROC) score of 0.9665. Wang et al [22] introduced a deep learning framework, Deep Learning for Drug-Drug Synergy prediction (DeepDDS), for predicting drug-drug interactions for anticancer treatments. DeepDDS uses gene expression data from the cancer cell line and the molecular graph of the drugs as input. It leverages GAT and graph convolution transformers (GCTs) to accurately predict the synergistic effect between drug combinations. DeepDDS has achieved an AUROC score of 0.67 on an independent test set. In the task of ADRD prediction, GCT obtained an area under the precision-recall curve of 0.34 on the inpatient and outpatient electronic health record (EHR) data from NYU Langone Health (briefly called AD-EHR) [23]. Klepl et al [24] integrated functional connectivity methods with GNNs to evaluate ADRD prediction performance using electroencephalography brain data. They showed that the GNN-based approach outperformed convolutional neural network and support vector machine models and obtained an AUROC of 0.984 [24]. Zhu and Razavian [23] presented variational GNN (VGNN), a variationally regularized encoder-decoder GNN, designed specifically for EHRs. This framework showed robustness in learning graph structures by applying regularization techniques to node representations. VGNN was used for ADRD risk prediction, and it attained an area under the precision-recall curve of 0.46 when using AD-EHR.

The abovementioned GNN models [23,24] have demonstrated the potential to uncover hidden patterns, reveal biological insights, and facilitate advancements in ADRD prediction. However, because the GNN architecture is a black-box model, the absence of interpretability is harmful to both users and society [25], especially in critical applications where decisions need to be explained or understood. Even though some advanced models such as GAT, GCT, and VGNN have the ability to explain the importance of individual nodes by using attention mechanisms, they still face a limitation in their interpretability concerning the significance of underlying relationships in the prediction process. As a consequence, there is a pressing demand for research and development efforts to enhance GNNs and elucidate the influence of relationship importance in achieving more precise ADRD predictions. By addressing this interpretability issue, GNNs can become more valuable tools in advancing our understanding of ADRD and contributing to improved patient care and treatment strategies.

Objective
The first focus of this study lies in the domain of risk prediction for ADRD. In this particular context, our investigation aims to
use claims data as the sole input for our GNN-based predictive model for accurate ADRD risk prediction. We enhance the predictive power of our model by incorporating advanced GNN models into a framework that enables us to effectively capture the intricate relationships and dependencies inherent in the claims data.

Second, we introduced a novel method to assess the importance of relationships within the patients’ individual medical record graphs and their influence on ADRD risk prediction. Generally, an additional graph explanation technique, such as GNNExplainer [26], is used as a post hoc method to interpret the predictions made by the GNNs. However, our proposed relation importance method enables an “in-process” explanation approach that leverages the relation weights from each patient’s individual graph. This method facilitates the interpretation of the GNN’s predictions during the graph generation process itself. Besides that, our method aims to adequately calibrate the importance of each relationship within the graph, reflecting their true impact on prediction. Since, typically, when a relation connects to nodes that are highly prevalent in the graph, there is a risk of misdefining its significance. The frequent occurrence of these nodes can distort the perception of the relationship’s importance, potentially leading to erroneous interpretations or biased conclusions. This bias can result in a skewed importance assigned to relationships, and hence potentially affecting the accuracy of ADRD risk prediction.

Methods

Cohort Description

We used deidentified administrative health claims data from Optum’s Clininformatics Data Mart, spanning from 2007 to 2020. This data set comprises over 68 million patient-level enrollment records submitted by various health care providers, pharmacies, and other health care service organizations for reimbursement purposes. It is accessible for researchers through a subscription provided by the University of Texas Health Science Center (UTHealth) [27].

Several criteria were applied to construct the study cohort, as illustrated in Figure 1. Considering that ADRD primarily affects older individuals and is a chronic condition, we initially filtered out patients (n=62,903,997) who were younger than 65 years. To ensure a sufficient data history for tracking their medical conditions, patients (n=2,680,329) with a time span of less than 3 years between their initial and final medical records were excluded. Patients (n=321,462) who lacked demographic information were also excluded from the study. To further establish the ADRD cohort, we used the definition outlined by Kim et al [28]. Patients were classified as having ADRD if they presented specific diagnosis codes or were prescribed relevant medications. The specified diagnosis codes are Alzheimer dementia (331.0*/G30.*), vascular dementia (290.4*/F01.*), frontotemporal dementia (331.1*/G31.0*), lewy body dementia (331.82*/G31.83), senile dementia (290.0*), presenile dementia (290.0*), presenile dementia (290.1*), other specified senile psychotic (290.8*), and unspecified senile psychotic condition (290.9*), and the medication includes aricept, donepezil, razadyne, reminyl, galantamine, exelon, rivastigmine, namenda, memantine, acetylcholine, and memantine. Based on the criteria mentioned above, the resulting cohort included 432,374 patients with ADRD and 1,895,511 patients without ADRD.
Data Preprocessing

In this study, we used a partitioning approach to categorize each patient’s records into 3 time windows, such as an index selection window, a feature window, and a prediction window (shown in Figure 2). First, we designated a specific period before the initial diagnosis of patients with ADRD or the last record for patients without ADRD as the index selection window. In the real world, patients may seek consultations for their health conditions at any time. To simulate this visiting setting, we randomly select the index day within each patient’s index selection window instead of using a fixed day. The 3-year period before the index day serves as the feature window for model training purposes, while a certain period after the index day is defined as the “prediction window” for ADRD risk prediction. Additionally, we designed 3 scenarios with index selection windows and prediction windows of 1, 2, and 3 years in length, respectively. By using this partitioning approach, we can comprehensively evaluate our model’s predictive accuracy in dynamically predicting ADRD diagnoses. It should be noted that researchers can easily adjust the lengths of these windows to align with their specific requirements and objectives.

There are also other inclusion criteria that were applied to ensure the quality of the data and the fairness of the cohort. Specifically, within the feature window, it was required that each patient have a minimum of 2 month-level records. Furthermore, within the records in the same month, a minimum of 3 medical codes (eg, diagnosis codes, procedure codes, and medication codes) needed to be present. After applying these criteria, the resulting cohort for each scenario is presented in Figure 1. In scenario 1, the cohort consisted of a total of 2,031,320 patients, comprising 228,086 patients with ADRD and 1,803,234 patients without ADRD. For scenario 2, the cohort comprised 2,007,625 patients, including 225,757 patients with ADRD and 1,781,868 patients without ADRD. Finally, in scenario 3, the cohort encompassed 1,961,641 patients, with 221,816 patients with ADRD and 1,739,825 patients without ADRD. These cohorts provide a robust foundation for further analysis and investigation in this study.

The data used in all cohorts included claims data consisting of diagnoses encoded with both International Classification of Diseases, 9th revision (ICD-9) codes and 10th version (ICD-10) codes, the National Drug Code for pharmacy claims, current procedural terminology, and Healthcare Common Procedure Coding System codes for procedures. The inclusion of both ICD-10 and ICD-9 codes was necessary as the study period spanned the transition from ICD-9 to ICD-10 coding systems. All these different types of medical codes have been converted to a higher-level categorization scheme to achieve feature reduction, uniformity, and compatibility within the study analysis. The ICD-9 and ICD-10 codes and the current procedural terminology and Healthcare Common Procedure Coding System codes are converted to clinical classification software, which is a tool for clustering patient diagnoses and procedures into a manageable number of clinically meaningful categories developed at the Agency for Healthcare Research and Quality (formerly known as the Agency for Health Care Policy) [29]. Similarly, we are using the Pharmacologic-Therapeutic Classification System from the American Hospital Formulary Service to represent and group the drug National Drug Code in the data set [30]. It is a method of grouping drugs with similar pharmacologic, therapeutic, and
chemical characteristics in a 4-tier hierarchy associated with a numeric code consisting of 2 to 8 digits. By following the conversion of these codes, the number of features was reduced from tens of thousands to hundreds. This reduction not only helps address the issue of sparsity in the model input but also improves its overall efficiency.

**Figure 2.** The definition of 3 scenarios. We established a time frame that includes an index selection window, a feature window, and a prediction window. The index selection window spanned a specific period before the initial diagnosis of patients with Alzheimer disease and related dementias (ADRD) or the last record for patients without ADRD. We randomly selected a day within the index selection window as the index day to simulate real-world visiting settings. The period up to 3 years before this index day was considered the feature window for training the model, while the period after the index day was used as the prediction window. We used 1 year, 2 years, and 3 years as the lengths of the index selection window and corresponding prediction window, respectively, to predict ADRD diagnosis dynamically.

**Modeling**

We used the VGNN in combination with patients’ diagnosis, procedure, and medication codes to estimate the likelihood of patients having ADRD within a designated prediction window. VGNN consists of 4 modules, such as the encoder graph, variational regularization, decoder graph, and fully connected layer. In the encoder graph module, VGNN takes 3 types of patients’ medical codes from the feature window as input and constructs a fully connected graph comprising medical codes for each patient. The representation of each node is iteratively updated through multiple graph attention layers. To address the challenges of generating node embeddings within clusters and achieving balanced attention weights, VGNN incorporates a variational regularization layer. This layer helps prevent model collapse and maintains the model’s expressive capacity. The decoder graph module uses the node representations generated by the encoder graph and the variational regularization layer to compute the weighted relations between each node. These weighted relationships effectively capture the relationships among different medical codes. Finally, a linear feed-forward layer is used to calculate the probability and produce the binary classification for identifying an individual with ADRD.

We initiated the modeling process by reserving 20% of patients from the entire data set for testing purposes. Given that ADRD is more prevalent in the older population [1] and our data set exhibits a high imbalance, we used the propensity score matching method based on age and gender to mitigate potential biases associated with these factors. This matching process ensured that our model’s input cohort consisted of individuals with similar age and gender distributions, reducing the potential confounding effects associated with these variables. As a result, we created a balanced cohort with a one-to-one ratio of control and case groups from the remaining 80% of the entire data set. This downsampling approach is a popular method in clinical research to create a balanced covariate distribution between treated and untreated groups, which could help significantly improve the model’s ability to handle imbalanced data [31]. We named it the matched cohort and used it for the purposes
of model development and validation. Additionally, we generated a smaller subset named the subset cohort, which is 10% of the matched cohort. This action allows us to evaluate the model’s performance on a smaller-scale data set effectively. In order to assess the efficacy of our approach, we built models for 3 different scenarios. Moreover, we used random forest (RF) and light gradient boost machine (LGBM) as baseline models and compared their performance with that of VGNN. The overall workflow of our model pipeline is shown in Figure 3.

Figure 3. The workflow of our study pipeline, including data preprocessing, graph modeling, and final output. The DIAG, PROC, and DRUG represent 3 types of medical codes: diagnosis, procedure, and medication, respectively. We used the variationally regularized encoder-decoder graph neural network (VGNN) to predict the likelihood of Alzheimer disease and related dementias (ADRD) using patients’ medical records sourced from Optum Clininformatics. The data were input into the encoder layer of VGNN, generating a fully connected graph specific to each patient. The variational regularization layer was then applied to prevent issues like mode collapse and maintain the model’s capacity to represent information effectively. Additionally, the decoder graph module used node representations to compute weighted relations between nodes, which effectively captured relationships among different medical codes. Finally, a linear feed-forward layer was used to calculate probabilities and perform binary classification.

Relation Importance
After the completion of model training, we then used the trained model to build the interconnected medical record graph for each individual patient. In order to evaluate the significance of various relationships in ADRD prediction, we extracted adjacency matrices from the medical graphs of N patients in the training set of the matched cohort. The values within these adjacency matrices serve as indicators of the relational importance associated with predicting ADRD. Given that the generated graphs are directional, the adjacency matrices \( A \) are not symmetric. Therefore, we took an additional step to mitigate the influence of directionality by computing the average of the original adjacency matrix and its transposed matrix. Then, the updated adjacency matrix is:

\[
\text{This adjacency matrix enables us to gain insights into the intricate relationships between medical codes and their predictive power for ADRD.}
\]

However, it is crucial to consider that medical codes with higher frequencies may have received relatively larger weights compared to others, potentially introducing bias in the analysis. Given that \( A^+ \) are the adjacency matrices of patients with ADRD case group and \( A^- \) are the adjacency matrices of ADRD patient control group, we calculated the mean adjacency matrix of these 2 patient groups as:

By subtracting the negative mean adjacency matrix from the positive mean adjacency matrix, we eventually obtained a mean weight-difference matrix:

This mean weight-difference matrix \( W \) captured the relative significance of the medical code weights. A higher positive value inside \( W \) indicates a greater importance in predicting ADRD, while a lower negative value suggests a reduced likelihood of ADRD occurrence. A value of 0 in \( W \) means that the relationship does not affect a patient with ADRD.

Ethical Considerations
The approval for the use of data in this study was obtained from the UTHealth Committee for the Protection of Human Subjects, under protocol HSC-SBMI-21-0965, with a waiver of consent granted.

Results

Hyperparameter Setting
We trained the VGNN model with the following hyperparameters: a learning rate of 0.0001, a batch size of 128, and a dropout rate of 0.1. We used the Adam optimizer for gradient descent and trained the model for 200 epochs. The model consisted of 2 graph layers and 1 attention head. To balance the binary cross-entropy and Kullback-Leibler divergence losses, a parameter value of 0.002 was used. Additionally, edge information was extracted after the attention layer to facilitate future calculations of relational importance. Additionally, we used the grid search method to tune the RF and LGBM baseline models. The hyperparameters for RF and
LGBM are \( n_{\text{estimators}}=100, \) \( \text{min}_{\text{samples}}_{\text{split}}=2, \) and \( \text{min}_{\text{samples}}_{\text{leaf}}=1, \) and \( n_{\text{estimators}}=300, \) \( \text{boosting}_{\text{type}}=\text{"gbdt,"} \) \( \text{num}_{\text{leaves}}=31, \) and \( \text{learning}_{\text{rate}}=0.1, \) respectively.

**Performance Evaluation**

We used AUROC as a measurement to evaluate the performance of each model. As shown in Table 1, the VGNN model achieved AUROC scores of 0.7272 and 0.7480 for the subset cohort and the matched cohort, respectively, in scenario 1. It outperformed the RF and LGBM models by an average of 10.6% and 9.1% across the 2 data sets. For scenario 2, the VGNN model obtained AUROC scores of 0.7125 and 0.7281 for the subset cohort and matched cohort, respectively. It exhibited superior performance compared to the other 2 models by an average of 10.5% and 8.9% across the 2 data sets. Finally, in scenario 3, the VGNN model achieved AUROC scores of 0.7001 and 0.7187, which were surpassed by an average of 10.1% and 8.5% across the 2 data sets. The results clearly demonstrate that the GNN approach (VGNN) outperforms the tree-based models (RF and LGBM) significantly in predicting ADRD. The bar chart for the performance comparison can be found in Multimedia Appendix 1.

Furthermore, we identified the 5 most important relationships for both positive and negative predictions of ADRD in Table 2. Among the top 5 negative highest-weighted relationships, “neoplasms of unspecified nature or uncertain behavior” exhibits its influence across all relations within scenario 1, “consultation, evaluation, and preventative care” makes a total of 4 appearances within scenario 2, while “quinolone antibiotics” spans all relations in scenario 3. Within the set of the top 5 positive highest-weighted relationships, both “routine chest x-ray” and “electrocardiogram” appear 3 times each in scenario 1, “substance-related disorders” contributes to 4 relationships in scenario 2, and “substance-related disorders” emerges as the most frequently occurring medical code in scenario 3.

*Table 1.* The model performance (area under the receiver operating characteristic curve scores) for Alzheimer disease and related dementias risk prediction.

<table>
<thead>
<tr>
<th>Scores</th>
<th>Random forest</th>
<th>Light gradient boost machine</th>
<th>Variational graph neural network</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Matched cohort</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 1</td>
<td>0.6710</td>
<td>0.6809</td>
<td>0.7480</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>0.6565</td>
<td>0.6658</td>
<td>0.7281</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>0.6468</td>
<td>0.6589</td>
<td>0.7187</td>
</tr>
<tr>
<td><strong>Subset cohort</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 1</td>
<td>0.6629</td>
<td>0.6720</td>
<td>0.7272</td>
</tr>
<tr>
<td>Scenario 2</td>
<td>0.6474</td>
<td>0.6570</td>
<td>0.7125</td>
</tr>
<tr>
<td>Scenario 3</td>
<td>0.6425</td>
<td>0.6490</td>
<td>0.7001</td>
</tr>
</tbody>
</table>
Table 2. Top 5 positive highest-weighted relations and top 5 negative highest-weighted relations.

<table>
<thead>
<tr>
<th>Scenarios and relations</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Top 5 negative highest-weighted relations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Neoplasms of unspecified nature or uncertain behavior</td>
<td>Consultation, evaluation, and preventative care</td>
<td>Dihydropyridines</td>
</tr>
<tr>
<td>2</td>
<td>Neoplasms of unspecified nature or uncertain behavior</td>
<td>Lens and cataract procedures</td>
<td>Consultation, evaluation, and preventative care</td>
</tr>
<tr>
<td>3</td>
<td>Neoplasms of unspecified nature or uncertain behavior</td>
<td>Hyperlipidemia</td>
<td>Consultation, evaluation, and preventative care</td>
</tr>
<tr>
<td>4</td>
<td>Neoplasms of unspecified nature or uncertain behavior</td>
<td>Diabetes mellitus with complications</td>
<td>Consultation, evaluation, and preventative care</td>
</tr>
<tr>
<td>5</td>
<td>Neoplasms of unspecified nature or uncertain behavior</td>
<td>Diagnostic ultrasound of head and neck</td>
<td>Diseases of white blood cells</td>
</tr>
</tbody>
</table>

| **Top 5 positive highest-weighted relations** | | | |
| 1 | Routine chest x-ray | Electrocardiogram | Substance-related disorders | Electrocardiogram | Schizophrenia and other psychotic disorder | Substance-related disorders |
| 2 | Routine chest x-ray | Other laboratory | Substance-related disorders | Other laboratory | Schizophrenia and other psychotic disorder | Diagnostic procedures on nose, mouth, and pharynx |
| 3 | Routine chest x-ray | Heart valve disorders | Substance-related disorders | Routine chest x-ray | Diagnostic ultrasound of head and neck | Arthrocentesis |
| 4 | Electrocardiogram | Other laboratory | Electrocardiogram | Inguinal and femoral hernia repair | Diagnostic ultrasound of head and neck | Substance-related disorders |
| 5 | Electrocardiogram | Heart valve disorders | Substance-related disorders | Coronary atherosclerosis and other heart disease | Substance-related disorders | Other diagnostic radiology and related techniques |

**Discussion**

**Principal Findings**

Based on this study’s results, we found that some potential candidates might be relevant to ADRD risk prediction and treatment. Our self-explainable GNN prediction method reveals the underneath connections between medical codes for ADRD risk prediction. Some code pairs have been shown to accelerate ADRD progression, while others exhibit potential to slow down its development. When implementing our relation importance interpretation method, the GNN results are explainable, setting it apart from other deep learning models. Moreover, several code pairs extracted from the GNN align with findings from previous research. Those code pairs that are not proven could offer valuable insights beyond the scope of current studies. In the following sections, we will present examples of code pairs derived from the GNN model results and highlight their significance based on validated evidence from previous studies.

This study found that certain pairs of medical codes can be associated with a decreased likelihood of an ADRD diagnosis. For instance, the treatment of more acute conditions, such as cancer or neoplasms, may delay the diagnosis of ADRD. We hypothesize that “neoplasms of unspecified nature or uncertain behavior” may be correlated with higher health care use or more frequent physician visits, similar to the code “consultation, evaluation, and preventative care.” The cooccurrence of these 2 types of coding could potentially lower the risk of ADRD. Regular health care visits could potentially reduce the risk of ADRD by improving modifiable risk factors and mitigating social isolation in older patients. Lee et al [32] revealed that cataract extraction is linked to a reduced risk of developing dementia among older adults. Cataract extraction has been associated with enhanced engagement in intellectually stimulating activities, such as reading and video consumption, as well as increased physical activity. These changes in lifestyle and cognitive engagement following cataract surgery may
contribute to a delay in the onset of ADRD. Consequently, the second node pair involving “neoplasms of unspecified nature or uncertain behavior” and “lens and cataract procedures” also holds relevance and supports the observed association. In scenario 2, Peters et al [33] have indicated that the use of calcium channel blockers, specifically dihydropyridines, is associated with a lower decline in cognitive function compared to other hypertensive treatments. Thus, the presence of the “consultation, evaluation, and preventative care” and “dihydropyridines” nodes pair ranking first in importance is consistent with the reported associations. The most frequently appearing node in scenario 3 is “quinolone antibiotics.” According to the study by Pham et al [34], it is a class of medication commonly prescribed to treat various bacterial infections and is primarily used for its antimicrobial properties [34]. Additionally from a study by Gao et al [35], their review study indicates that the brain inflammation caused by microbial infections may be one of the etiologies of ADRD, and antibiotics as novel treatments may be beneficial for delaying the development of ADRD. Quinolones exhibit a distinct pharmacokinetic profile characterized by a higher cerebrospinal fluid to serum concentration ratio compared to other commonly prescribed antibiotics [36]. This unique attribute may underlie the observed robust negative correlation between quinolone administration and the development of ADRD, distinguishing its potential protective effect from that of other antibiotics. The use of quinolones likely correlates with younger age, as its use in older adults is less frequent due to the increased risk of tendon rupture. However, this is less likely to explain its negative correlation with the onset of ADRD in our age-matched cohorts. So, in other words, it can be hypothesized that “quinolone antibiotics” may potentially exhibit a slowing effect on the progression of ADRD. Combined with the aforementioned node “lens and cataract procedures,” the observed association of this node pair holds validity and is worth further investigation.

This study also found certain medical codes to be positively associated with a higher likelihood of an ADRD diagnosis. This can be explained by the fact that Alzheimer disease, to a certain degree, is a “diagnosis of exclusion.” Procedures like “routine chest x-ray” and “electrocardiogram” are commonly used as initial steps in diagnosing altered mental status, which is often the first sign of ADRD. A chest x-ray is often used to rule out any underlying pneumonia, while an electrocardiogram may be used to rule out arrhythmia [37]. Similarly, “diagnostic ultrasound of head and neck” is commonly done to rule out conditions like carotid artery clots, stenosis, or plaque in the setting of stroke workups. Once patients begin to verify these initial diagnoses of altered mental status, they are more likely to undergo comprehensive and relevant testing to exclude other potential causes of the symptoms, which may potentially lead to a timely determination of ADRD. Several studies have also found that alcohol and drug use could affect mental state and cognitive function [38]. People who abuse intoxicating substances for a considerable period may develop dementia or accelerate the neurological damage associated with Alzheimer [39].

From the modeling aspect, to the best of our knowledge, our approach offers distinct advantages in comparison to previous studies on the early diagnosis of ADRD with or without GNN methods. For instance, Li et al [40] used a gradient boost tree and logistic regression to assess ADRD risk using EHR data from the OneFlorida+ Clinical Research Consortium. They identified significant clinical and social factors through SHAP values; however, these factors were commonly known risk factors. In contrast, our findings unveil potential risk factors and explain the interaction among these factors in ADRD prediction. While VGNN demonstrates good interpretability by showcasing attention weights among features, it fails to explain how these features positively or negatively impact ADRD prediction [23]. On the other hand, our model offers interpretations of potential risk factors and illustrates their influence on outcomes. Furthermore, our proposed self-explainable framework mitigates the potential bias resulting from the prevalence of medical codes. Klepl et al [24] conducted electroencephalography-based ADRD prediction using GNN methods. As medical image data are unavailable for every patient during routine examinations, limitations arise due to the restricted user cases. Furthermore, they only assessed model performance against baseline models without providing any feature interpretation. Conversely, our method enhances interpretability by leveraging more accessible data, thereby promoting broader applicability and understanding in ADRD prediction. Overall, this is the first work that proposes a self-explainable framework, providing a feature-important explanation in the context of ADRD risk prediction leveraging relational information within a graph. Compared with other studies on ADRD risk predictions, our method can directly interpret the relationship’s importance within the training process. It does not require any additional post hoc explanation methods, such as GNNExplainer [26]. In other words, within our framework, it takes no additional time to get an explanation.

In summary, we showed that using the GNN approach for ADRD prediction has better performance compared to baseline models. Moreover, with the incorporation of our relation importance method, the model’s results become explainable, providing valuable insights into the underlying factors contributing to ADRD risk prediction.

Limitations

Our prediction does not incorporate time information into the modeling process. In this study, we aggregated 3 years of records into a single representation and treated them equally without considering their temporal sequence. In the real-world clinical setting, medical events, procedures, or medications obtained at different times should carry different levels of significance. In other words, events occurring closer in time to the prediction window are expected to have a greater impact on the disease prediction. In our future study, we could use a time series model and positional encoding method to establish connections between patients’ multiple visit records for more accurate predictions and provide more valuable insights into ADRD prediction.

On the other hand, it is important to note that certain predicted correlations may not causally assist clinicians in diagnosing ADRD. For instance, initiating tests for early detection of altered mental status might lead patients to identify ADRD through
various related tests. Nonetheless, from the clinician’s perspective, ordering these test results may not be helpful for early ADRD prediction. In our future work, we could try to exclude these “subjective patient-related factors” and instead focus on investigating more objective risk factors that could potentially influence the prediction of ADRD.

Conclusion
In this study, we used an advanced self-explainable GNN approach and developed a relation importance interpretation method for the ADRD risk prediction task based on claims data sources. The VGNN model’s effectiveness was evaluated across 3 distinct scenarios, with comparisons made against RF and LGBM ML models. The model’s performance achieved satisfactory results. In addition, we provided the interpretation for the node pairs extracted from the KG, which was generated from the VGNN model. Furthermore, we demonstrated the results’ future applicability and explained the important node pairs that align with previous research findings. This work contributes to the advancement of ADRD prediction models and reinforces the importance of interpretable results for informed clinical decision-making and early detection, etc.

Acknowledgments
We thank Luyao Chen and Xiaoqian Jiang for extracting data from Optum’s Clinformatics. This research was partially supported by National Institutes of Health grants under award numbers R01AG084236, R01AG083039, RF1AG072799, and R56AG074604.

Authors’ Contributions
CT conceived the research project, supervised the research, and critically revised the manuscript. XH, ZS, and YN designed the pipeline and method and implemented the deep learning model of the study and the explanation method. XH, ZS, and YW prepared the manuscript. XH and YD provided the cohort selection, scenario definitions, model workflow, and model performance figures. XH and JF prepared the data. FL and EY provided suggestions on data filtering, index day selection, and model design. YW composed a clinical discussion on the principal findings. All authors proofread the paper, provided valuable suggestions, and have read and approved the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
The model performance (area under the receiver operating characteristic scores) for Alzheimer disease and related dementia risk prediction. Matched cohort: employ 1:1 propensity score match for case and control in the original training data (i.e., 80% of full data), to achieve a balanced dataset and train the models, and test the models in the hold-out 20% of full data. Subset cohort: use around 10% of the Matched cohort (i.e., 20,000 for both case and control) to train the models, and test the models in the hold-out 20% of full data.

References
27. Optum's Clininformatics Data Mart. The University of Texas Health Science Center at Houston. URL: https://sbmi.uth.edu/sbmi-data-service/data-set/optum/ [accessed 2022-06-06]


Abbreviations

ADRD: Alzheimer disease and related dementias
AUROC: area under the receiver operating characteristic curve
DeepDDS: Deep Learning for Drug-Drug Synergy prediction
EHR: electronic health record
GAT: graph attention network
GCT: graph convolution transformer
GNN: graph neural network
ICD-9: International Classification of Diseases, 9th revision
ICD-10: International Classification of Diseases, 10th revision
LGBM: light gradient boost machine
ML: machine learning
RF: random forest
VGNN: Variationally regularized encoder-Decoder Graph Neural Network

©Xinyue Hu, Zenan Sun, Yi Nian, Yichen Wang, Yifang Dang, Fang Li, Jingna Feng, Evan Yu, Cui Tao. Originally published in JMIR Aging (https://aging.jmir.org), 08.07.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Advocating for Older Adults in the Age of Social Media: Strategies to Achieve Peak Engagement on Twitter

Reuben Ng¹², PhD; Nicole Indran¹, BSocSci (Hons); Luyao Liu², MSc

Corresponding Author:
Reuben Ng, PhD

Abstract

Background: Over the last decade, many organizations dedicated to serving the needs and interests of older adults have turned to social media platforms, such as Twitter, subsequently rebranded X, to improve the visibility of age-related issues. However, notwithstanding their growing digital presence and participation, minimal attention has been paid to the use of social media among these advocacy groups. To achieve policy change, advocacy organizations must first be able to engage and mobilize audiences.

Objective: Our study aims to elucidate how different tweet features affect the time it takes for posts uploaded by age advocacy organizations to reach peak engagement.

Methods: We collated 204,905 tweets from 53 age advocacy organizations posted over a 12-year period. The engagement score of each tweet was calculated by combining well-established metrics, namely likes, retweets, quote tweets, and replies. We ran Cox models with tweet features as predictors and time-to-peak engagement as the outcome. “Peak engagement” (event) refers to engagement scores above the 75th percentile, and “time” refers to months taken to reach peak engagement per tweet.

Results: Approximately 1 in 2 tweets (n=103,068, 50.3%) had either no hashtags or just 1 hashtag. Around two-thirds (n=131,220, 64%) of the tweets included a URL. Visual information was highly underused, with most tweets not including GIFs (n=204,202, 99.7%), videos (n=199,800, 97.5%), or photos (n=143,844, 70.2%). Roughly half (n=101,470, 49.5%) of the tweets contained mentions and 9.3% (n=19,009) of tweets were replies. Only 4.5% (n=9285) of tweets were quote tweets. Most tweets were uploaded in the afternoon (n=86,004, 42%) and on a weekday (n=180,499, 88.1%). As hypothesized, features associated with peak engagement were the inclusion of visual elements like photos, which increased peak engagement by 3 times (P<.001), and the use of 3 or more hashtags (P<.001). Quote tweets increased engagement by 3 times (P<.001), as compared to regular tweets, controlling for account-level covariates. Tweets from organizations with a higher tweet volume were 40% less likely to reach peak engagement (P<.001).

Conclusions: Social media as a networked platform has the potential to reach users on a global scale and at an exponential speed. Having uncovered the features that are more likely to reach peak engagement on Twitter, our study serves as an invaluable resource for age advocacy organizations in their movement to create a more age-inclusive world.

(JMIR Aging 2024;7:e49608) doi:10.2196/49608

KEYWORDS

age advocacy; social media engagement; older adults; ageism; data science

Introduction

Over the last decade, many organizations dedicated to serving the needs and interests of older adults have turned to social media platforms, such as Twitter, consequently rebranded as X, to improve the visibility of age-related issues. However, notwithstanding their growing digital presence and participation, minimal attention has been paid to the use of social media among these advocacy groups. To achieve policy change, advocacy organizations must first be able to engage and mobilize audiences. Our study elucidates how different tweet features affect the time it takes for posts uploaded by age advocacy organizations to reach peak engagement. We define “age advocacy” as the act of supporting or championing initiatives that address the needs of older adults.

The advent of social media has been profitable to advocacy groups for multiple reasons. First, although the visibility of a social movement was formerly determined by its ability to make headlines [1], social media has democratized the process of activism, allowing social actors to bypass the lack of attention received by a particular issue [2]. Second, the exchange of information via social media channels is not constrained by geographical barriers, thus enabling the rapid diffusion of information worldwide [1]. Third, social media platforms are a cost-effective means through which information can be transmitted and awareness of social issues heightened. Fourth,
social media facilitates interaction between organizations and the public, thereby fostering sociopolitical discussion and participation [3].

When using social media, organizations typically set out to engage followers by uploading content that resonates with audiences [4] especially in view of the constant influx of information on the internet [5,6]. A well-engaged audience is essentially proof that a particular account has content which audiences find valuable and meaningful. Over the years, this concept of engagement has gained popularity across myriad disciplines, including marketing, psychology, communication, public relations, and organizational studies [3].

Twitter is a microblogging service home to over 300 million active users monthly [7]. Although originally viewed as an avenue for personal communication, the social media platform has since been used by academics, policy makers, and advocacy groups to access, share, and disseminate information [8]. Given the growing presence of age advocacy organizations on Twitter, this study looks at how different tweet features affect the time taken to reach peak engagement for posts uploaded by these organizations.

Both marketing experts and academics have conducted research on the features that promote user engagement on Twitter [9-15]. Although it is clear that adding photos and videos improves engagement [10,14], it remains a scholarly crux when the best time to post is [9,10,14,15], what the ideal number of hashtags to include is [11,12,14], and whether quote tweets drive engagement. There is, therefore, a need to ascertain which tweet features are linked to greater user engagement for content uploaded by age advocacy organizations specifically.

To date, only 1 study has explored the concept of engagement in relation to tweets uploaded by age advocacy organizations [16]. However, this study did not consider the time taken to reach peak engagement, which is important for several reasons. First, the time taken to hit peak engagement may be viewed by potential funders as a key performance indicator, which is a signal of the ability of an organization to retain the interest of its user base and consequently be eligible for further funding. Second, being able to reach peak engagement within a short period of time is vital if age advocacy organizations happen to be posting about time-sensitive issues.

From a conceptual angle, this study is significant in that it is one of the first to develop a framework that age advocacy organizations can use to optimize their social media posts for increased engagement. Existing studies have traced the origins of age advocacy in the United States [17] and have covered the need to advocate for older persons [18-24]. Research on web-based age advocacy, however, remains conspicuously absent, with most social media analyses in the gerontological field analyzing attitudes toward older persons [25-32]. From a practical angle, this study provides organizations with a road map to raise consciousness of age-related matters, which is especially pressing given the increasing proportion of older adults in populations worldwide [33]. By successfully engaging audiences, age advocacy organizations will be able to spur collective action and create policy change.

The tweet features examined in this study include the number of hashtags, URLs, and mentions present in the tweet; whether the tweet contains a GIF, photo, or video; whether the tweet is a “quote tweet”—a retweet with a comment added by the account—or a “reply”; the time of day the tweet was uploaded; and whether the tweet was uploaded on a weekday or the weekend.

We sought to test 4 hypotheses. First, in light of past findings that the inclusion of hashtags predicts the likelihood of a post to get retweeted [9,10,12], we hypothesized that tweets with more hashtags would be quicker to reach peak engagement (hypothesis 1). Second, in line with evidence that visual information is usually more stimulating than textual information [34,35], we hypothesized that tweets with GIFs, photos, or videos would be quicker to reach peak engagement than those without (hypothesis 2). Third, since followers of age advocacy organizations are likely to include scholars and policy makers who may value dialogue, input, or commentary [9,12,36,37], we hypothesized that quote tweets would be quicker to reach peak engagement (hypothesis 3). Finally, consistent with prior research, which finds higher tweet counts to be associated with negative engagement [9], we hypothesized that tweets uploaded by accounts with a higher tweet count would be slower to reach peak engagement (hypothesis 4).

Methods

Data Set

As few studies have looked at age advocacy organizations on Twitter, we first consolidated a list of organizations by referring to various sources [38-40]. Next, we checked whether these accounts had a presence on Twitter. To build a more comprehensive list of accounts, we looked through the list of followers of these accounts and identified other organizational accounts with large followings using a snowball sampling method. The organizations were eventually chosen based on the following inclusion criteria: (1) the organization was based in North America; (2) the organization was dedicated to serving the needs and interests of older persons specifically; and (3) the organization had at least 1000 followers. In total, there were 53 accounts (Multimedia Appendix 1).

We retrieved the data using the Twitter application programming interface (API) v2, which was accessed through Twitter’s Academic Research Product Track [41]. The v2 full-archive search allows for the programmatic access of public tweets from the complete archive dating back to the first tweet in March 2006, when the application was created. Relative to what was achievable with the standard v1.1 API, the v2 API grants users a higher monthly tweet cap and access to more precise filters [42].

Tweets collected (n=403,426) covered a period of 12 years, from July 17, 2009, to October 8, 2021, with the start date as the earliest date a particular tweet from any of the sampled accounts was uploaded and the end date a week after October 1, 2021, which was designated by the United Nations as the International Day of Older Persons [43]. “Retweets” (n=118,454) were excluded since they are not original content.
Similarly, tweets with zero engagement (n=80,065) were excluded, as our focus was to observe the time taken to reach peak engagement. Finally, due to glitches with the API during the period of data collection—there were inaccuracies in the number of “likes” received by certain tweets—a few posts (n=2) were excluded. The final data set comprised 204,905 tweets.

**Tweet Features (Predictors and Covariates)**

Similar to earlier work [9], we divided the tweet features into 2 categories: tweet-level (predictors) and account-level (covariates) features. The tweet-level features include the number of hashtags, URLs, and mentions present in the tweet; whether the tweet contains a GIF, photo or video; whether the tweet is a “quote tweet” or a “reply”; the time of day the tweet was uploaded; and the day—weekday or weekend—the tweet was uploaded. Following past literature [14], we divided the time of day based on CST into the following periods: morning (6 AM to 11:59 PM), afternoon (noon to 16:59 PM), evening (5 PM to 8:59 PM), and night (9 PM to 5:59 AM).

Account-level features, which served as covariates in our modeling, were consistent across all tweets belonging to a given account. These covariates included the number of followers, the number of accounts followed, the total number of tweets, and whether the account was “verified.” Except for the last variable, all skewed account-level variables were log transformed. Multimedia Appendix 2 contains a list of definitions of terms used on Twitter.

**Time-to-Peak Engagement (Outcome)**

Following Twitter’s data dictionary [44], we used “likes” (ie, the number of times a particular tweet has been liked by other Twitter users), “retweets” (ie, the number of times a particular tweet has been retweeted), “quote tweets” (ie, the number of times a particular tweet has been quoted by other Twitter users) and “replies” (ie, the number of times a particular tweet has been replied to) as a proxy for user engagement. Our measurement of engagement aligns with that of previous studies [9,10,12-14]. To model the temporal aspects of engagement, we applied methods from survival analysis [45,46], which involved operationalizing engagement as a time-to-event variable. “Peak engagement” (event) refers to engagement scores above the 75th percentile, and “time” refers to months taken to reach peak engagement per tweet.

**Analytic Strategy**

First, we performed Kaplan-Meier analyses to assess differences in engagement between categorical features—type of tweet and presence or absence of visual elements, such as photos, GIFs, videos, hashtags, URLs, and mentions. Respective curves were compared using the log-rank statistic. Second, we ran Cox regression models to identify the tweet features significantly associated with time-to-peak engagement, controlling for account-level variables. Since tweets from the same account contained identical account-level information, the independent assumption did not hold. To achieve a more robust variance, we set different user IDs as clusters [47]. Model 1 consisted of tweet-level features. Model 2 contained tweet-level features, controlling for account-level variables as covariates.

**Ethical Considerations**

Ethical approval was not deemed necessary, as all the data used were publicly available and anonymized.

**Results**

**Descriptive Statistics**

Approximately 1 in 2 tweets (n=103,068, 50.3%) had either no hashtags or just 1 hashtag. Around two-thirds (n=131,220, 64%) of the tweets included a URL. Visual information was highly underused, with most tweets not including GIFs (n=204,202, 99.7%), videos (n=199,800, 97.5%), or photos (n=143,844, 70.2%). Roughly half (n=101,470, 49.5%) of the tweets contained mentions, and 9.3% (n=19,009) of the tweets were replies. Only 4.5% (n=9285) of the tweets were quote tweets. Most tweets were uploaded in the afternoon (n=86,004, 42%) and on a weekday (n=180,499, 88.1%). Table 1 summarizes the descriptive statistics. With regard to engagement, the lowest score was 1, and the highest score was 18,558. The engagement score at the 75th percentile was 8. Of the 204,905 tweets, 48,103 received an engagement score above 8.
Table. Description of tweets (n=204,905) from 53 age advocacy organizations posted over 12 years.

<table>
<thead>
<tr>
<th>Tweet-level variables</th>
<th>Values, n (%)^a</th>
<th>F^b</th>
<th>P value^c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of hashtags</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 or 1</td>
<td>103,068 (50.3)</td>
<td>859.3</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>2</td>
<td>53,336 (26.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥3</td>
<td>48,501 (23.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of URLs</td>
<td></td>
<td>137.5</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>0</td>
<td>61,346 (29.9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>131,220 (64.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥2</td>
<td>12,339 (6.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of mentions</td>
<td></td>
<td>1252</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>0</td>
<td>103,435 (50.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>65,869 (32.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥2</td>
<td>35,601 (17.4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GIF</td>
<td></td>
<td>701.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No</td>
<td>204,202 (99.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>703 (0.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photo</td>
<td></td>
<td>11,540</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No</td>
<td>143,844 (70.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>61,061 (29.8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Video</td>
<td></td>
<td>894.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>No</td>
<td>199,800 (97.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>5105 (2.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of tweet</td>
<td></td>
<td>4800</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Original tweet</td>
<td>176,611 (86.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quote tweet</td>
<td>9285 (4.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reply</td>
<td>19,009 (9.3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of upload</td>
<td></td>
<td>498.2</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Afternoon</td>
<td>86,004 (42)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Evening</td>
<td>28,606 (14.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morning</td>
<td>81,041 (39.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Night</td>
<td>9254 (4.5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day of upload</td>
<td></td>
<td>132.1</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Weekday</td>
<td>180,499 (88.1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekend</td>
<td>24,406 (11.9)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

^aPercentages may not add up to 100 due to rounding.
^bF refers to the F-statistic for the ANOVA test.
^cP values are for the ANOVA test.

Kaplan-Meier Analysis: Differences in Engagement Across Tweet Features

We performed Kaplan-Meier analyses to examine differences in engagement across tweet features for 204,905 tweets posted over 146 months. Quote tweets achieved median engagement twice (log-rank test: $\chi^2=3820; P<.0001$) as fast as regular tweets (Figure 1). Specifically, there was an engagement advantage of 65 months, meaning that on average, quote tweets achieved peak engagement 65 months faster than regular tweets. Regarding visual elements, tweets with photos reached 75th percentile engagement 2.5 times faster than tweets without photos (log-rank test: $\chi^2=1070; P<.0001$), having an engagement advantage of 80 months (Figure 2). Similar results were
observed for tweets containing GIFs (log-rank test: $\chi^2=1070; P<0.0001$) and videos (log-rank test: $\chi^2=8069; P<0.0001$) as compared to tweets without the respective features. Tweets with 3 or more hashtags had an engagement advantage of 14 months as compared to those with 2 hashtags (log-rank test: $\chi^2=2700; P<0.0001$). Similar patterns emerged for URLs and mentions. Tweets with 2 or more URLs achieved an engagement advantage of 14 months compared to tweets with 1 URL (log-rank test: $\chi^2=514; P<0.0001$). Conversely, tweets without mentions had greater engagement advantage than tweets with at least 1 mention (log-rank test: $\chi^2=850; P<0.0001$).

**Figure 1.** Engagement trajectories for 204,905 quote tweets, regular tweets, and replies posted over 12 years.
Multivariable Cox Regression: Tweet Features Associated With Time-to-Peak-Engagement

Tweets with 3 or more hashtags were 75% more likely to reach peak engagement than those with 1 or no hashtags (hazard ratio 1.75; \( P < .001 \)), supporting hypothesis 1. Visual elements were particularly effective in nudging tweets toward peak engagement, being 4.25 times more effective for tweets with photos (\( P < .001 \)), 6.38 times more effective for tweets with GIFs (\( P < .001 \)), and 9.97 times more effective for tweets with videos (\( P < .001 \)). This provided support for hypothesis 2. Consistent with hypothesis 3, quote tweets were 3.15 times more likely to achieve peak engagement (\( P < .001 \)), as compared to regular tweets, controlling for account-level variables, such as the number of followers, the number of people followed, the number of tweets, and verified status. Meanwhile, at the account level, tweets posted by organizations with a higher tweet count were 40% less likely to reach peak engagement (hazard ratio 0.60; \( P < .001 \)) as compared to those with a lower tweet count, providing support for hypothesis 4. The regression results are presented in Table 2. Coefficients reached significance at \( P < .05 \) after correcting for multiple comparisons using the Bonferroni method [48].
Table. Multivariable Cox models of tweet-level and account-level predictors of time-to-peak engagement for tweets (n=204,905) posted by age advocacy organizations over 12 years. Variables were log transformed.

<table>
<thead>
<tr>
<th>Tweet-level variables</th>
<th>Model 1(^a)</th>
<th></th>
<th></th>
<th>Model 2(^a)</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Hazard ratio</td>
<td>P value(^b)</td>
<td>Hazard ratio</td>
<td>P value(^b)</td>
<td>Hazard ratio</td>
<td>P value(^b)</td>
</tr>
<tr>
<td></td>
<td>(95% CI)</td>
<td></td>
<td>(95% CI)</td>
<td></td>
<td>(95% CI)</td>
<td></td>
</tr>
<tr>
<td><strong>Number of hashtags</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 or 1</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>2</td>
<td>1.19 (0.95-1.50)</td>
<td>&gt;.99</td>
<td>1.19 (1.00-1.42)</td>
<td>.90</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥3</td>
<td>1.50 (1.12-2.02)</td>
<td>.11</td>
<td>1.75 (1.37-2.22)</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of URLs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>1</td>
<td>1.18 (0.94-1.48)</td>
<td>&gt;.99</td>
<td>1.25 (1.07-1.45)</td>
<td>.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥2</td>
<td>1.43 (1.01-2.01)</td>
<td>.62</td>
<td>1.44 (1.05-1.98)</td>
<td>.42</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Number of mentions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>1</td>
<td>0.80 (0.67-0.97)</td>
<td>.32</td>
<td>0.84 (0.71-0.99)</td>
<td>.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥2</td>
<td>0.95 (0.71-1.26)</td>
<td>&gt;.99</td>
<td>1.03 (0.82-1.29)</td>
<td>&gt;.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>GIF</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>8.63 (6.33-11.77)</td>
<td>&lt;.001</td>
<td>6.38 (3.62-11.24)</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Photo</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.06 (2.76-5.96)</td>
<td>&lt;.001</td>
<td>4.25 (3.19-5.68)</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Video</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.39 (6.06-29.58)</td>
<td>&lt;.001</td>
<td>9.97 (3.36-29.65)</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Type of tweet</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Original</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Quote tweet</td>
<td>3.03 (1.96-4.68)</td>
<td>&lt;.001</td>
<td>3.15 (2.02-4.90)</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Replies</td>
<td>0.41 (0.14-1.21)</td>
<td>&gt;.99</td>
<td>0.28 (0.07-1.17)</td>
<td>&gt;.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Time of upload</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Afternoon</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Evening</td>
<td>1.21 (1.01-1.45)</td>
<td>.63</td>
<td>1.09 (0.92-1.29)</td>
<td>&gt;.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morning</td>
<td>1.02 (0.90-1.15)</td>
<td>&gt;.99</td>
<td>0.98 (0.86-1.12)</td>
<td>&gt;.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Night</td>
<td>1.10 (0.81-1.48)</td>
<td>&gt;.99</td>
<td>1.07 (0.73-1.57)</td>
<td>&gt;.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Day of upload</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekday</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Weekend</td>
<td>1.00 (0.89-1.12)</td>
<td>&gt;.99</td>
<td>0.98 (0.88-1.08)</td>
<td>&gt;.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Account-level variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Follower count</td>
<td>—(^c)</td>
<td>—</td>
<td>1.42 (1.17-1.73)</td>
<td>.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friend count</td>
<td>—</td>
<td>—</td>
<td>1.60 (1.16-2.21)</td>
<td>.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tweet count</td>
<td>—</td>
<td>—</td>
<td>0.60 (0.49-0.73)</td>
<td>&lt;.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verified status</td>
<td>—</td>
<td>—</td>
<td>1.20 (0.78-1.86)</td>
<td>&gt;.99</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Constant not shown.
\(^b\)P values have been adjusted using Bonferroni correction.
\(^c\)Not applicable.

**Discussion**

**Principal Findings**

Although the technological era has ushered in numerous opportunities for advocacy organizations, scant attention has been devoted to examining the use of social media as a tool for age advocacy. As social media can be instrumental in fostering policy change, we sought to fill this gap by examining how different tweet features influence engagement for tweets uploaded by age advocacy organizations. Findings indicate that...
tweets that are more likely to reach peak engagement are those that include 3 or more hashtags, contain visual elements, or are quote tweets. In contrast, tweets posted by organizations with a higher tweet count are less likely to reach peak engagement as compared to those with a lower tweet count.

Unsurprisingly, tweets with hashtags are more likely to achieve peak engagement. Arguably the most iconic feature of Twitter is the hashtag, which is an organizational device that connects users to a broader community of individuals who use the same hashtag [1]. Although movements concerning race- or gender-related matters are notably associated with hashtags, such as #BlackLivesMatter and #MeToo, hashtags related to age have not garnered the same level of success. Nevertheless, several age advocacy groups have rolled out their own hashtag campaigns in recent years. For example, the American Association of Retired Persons started the hashtag #DisruptAging as a way to spark conversations on what it means to grow older [49]. Age Platform Europe, a network lobbying for the rights of older adults, began the #AgeingEqual campaign in 2018 to raise awareness of ageism [50]. More recently, the World Health Organization started the hashtag #AWorld4AllAges in a bid to encourage individuals to build a more age-inclusive world [51]. Moving forward, age advocacy organizations could consider embedding their tweets with more hashtags to improve the visibility of their content.

Peak engagement is also achieved when visual elements, such as GIFs, photos, or videos, are included in a tweet. That the brain absorbs and synthesizes visual information faster than textual information is an insight from past research [34,35]. Our results reveal that GIFs, photos, and videos are all piteously underused in content uploaded by age advocacy organizations. These organizations should therefore strive to include visual elements in their tweets to bolster their chances of capturing the attention of followers. Importantly, these elements should be carefully selected to avoid perpetuating visual ageism [52]. Organizations could consider selecting images from the newly launched Age-Positive Image Library, which houses images that portray old age more realistically [53].

Quote tweets reach peak engagement faster than original tweets. Whereas the retweet function enables users to repost a tweet verbatim, quote tweets give users the option of adding their own comments to the tweet being reposted and is often used by individuals who wish to express their opinions in the context of the original tweet [54]. Given how a large subset of those following age advocacy organizations likely comprises academics and policy makers—people who may rely on Twitter for sharing knowledge or participating in intellectual discussions [9,12]—it makes sense that quote tweets take less time to reach peak engagement. With less than 5% of the tweets collected being quote tweets, age advocacy organizations should consider using the quote tweet function more regularly to establish a dialogic relationship with the public.

As expected, having a higher follower count lessens the time needed to reach peak engagement. Both older and newer accounts should therefore make concerted efforts to amass as many followers as possible. In particular, age advocacy organizations with little or no digital presence should prioritize crafting strategies to increase their follower count before attempting to bolster engagement.

Not spamming audiences with content is considered by marketing experts to be a basic rule of Twitter etiquette [55]. By posting too often, organizations risk losing public interest or frustrating followers [9]. In seeking to forge a connection with the public, organizations must exercise prudence with regard to how frequently they post to prevent inundating followers’ feeds. There are no hard and fast rules about how often to tweet, but social media managers of age advocacy organizations could monitor levels of engagement using the platform’s “Tweet Activity Dashboard” [55]. By tracking the level of engagement of each tweet, organizations will be able to gain insight into the optimal frequency for tweeting.

As age advocacy organizations curate their content with the goal of maximizing engagement, it is imperative that these organizations extend their outreach beyond researchers and policy makers to the larger society. This is especially critical since age-related issues have yet to gain widespread awareness among the public. Moreover, age advocacy organizations could involve older adults in the cocreation of initiatives, such as by collaborating with older influencers [56,57]. In addition, amid the prevalence of intergenerational tension in the digital sphere [29,30], there is a need to create opportunities for older and younger generations to interact. Hashtag campaigns could be used to encourage both generations to engage in meaningful dialogues.

Limitations

This study has a number of limitations. First, the period that the tweets were posted is likely to have been a confounder in our analysis. It was only in 2014 that GIFs could be shared on Twitter. Likewise, the quote tweet feature was introduced only in 2015. However, tweets uploaded from 2009 onwards were included in our data set. The fact that there are now many more users on Twitter also means that posts that were uploaded before the platform was popular were less likely to be well engaged with. Second, considering that our objective was to look specifically at organizations, we could not offer insight into the level of engagement of tweets belonging to influential activists who champion the rights of older persons. Third, age advocacy organizations that are newer to Twitter were not included in the study since they did not fulfill the inclusion criterion of having at least 1000 followers at the time of analysis. Fourth, it is important to highlight that some tweets may have been uploaded solely for the purpose of informing or educating the public, rather than with the goal of engagement [9]. Finally, whether or not digital engagement actually inspires real-world action remains a moot point. Future analyses could adopt survey-based techniques [58,59] to understand activists’ perceptions of digital activism and how it compares to traditional offline activism.

Despite these limitations, our study contributes to the field of gerontology by developing some practical guidelines for improving age advocacy efforts on Twitter. With research on this topic still at the outset, directions for future research are plentiful. Foremost among them is the need to construct a theoretical framework outlining the concept of age advocacy. Subsequent research could also explore how levels of
engagement vary across organizations specializing in areas like retirement, housing, or health care. Additionally, it would be worthwhile to dissect the profile of followers of age advocacy organizations. This could include an analysis of the distribution of followers based on characteristics such as age, gender, and occupation.

Conclusions
Social media as a networked platform has the potential to reach users on a global scale and at an exponential speed. Having uncovered the features that are more likely to reach peak engagement on Twitter, our study serves as an invaluable resource for age advocacy organizations in their movement to create a more age-inclusive world.

Conflicts of Interest
None declared.

Multimedia Appendix 1
List of age advocacy organizations.
[DOCX File, 22 KB - aging_v7i1e49608_app1.docx ]

Multimedia Appendix 2
Definitions of terms used on Twitter.
[DOCX File, 20 KB - aging_v7i1e49608_app2.docx ]

References
41. Tornes A. Enabling the future of academic research with the Twitter API. Developer Platform. 2021 Jan 26. URL: https://tinyurl.com/bdthbscb [accessed 2021-10-13]


Abbreviations

API: application programming interface
Perceptions and Utilization of Online Peer Support Among Informal Dementia Caregivers: Survey Study

Zhijun Yin1,2, PhD; Lauren Stratton3, PhD; Qingyuan Song2, ME; Congning Ni2, ME; Lijun Song4, PhD; Patricia Commiskey5, DrPH; Qinxia Chen1,6, PhD; Monica Moreno7, BS; Sam Fazio8, PhD; Bradley Malin1,2,6, PhD

Corresponding Author:
Zhijun Yin, PhD

Abstract

Background: Informal dementia caregivers are those who care for a person living with dementia and do not receive payment (eg, family members, friends, or other unpaid caregivers). These informal caregivers are subject to substantial mental, physical, and financial burdens. Online communities enable these caregivers to exchange caregiving strategies and communicate experiences with other caregivers whom they generally do not know in real life. Research has demonstrated the benefits of peer support in online communities, but this research is limited, focusing merely on caregivers who are already online community users.

Objective: We aimed to investigate the perceptions and utilization of online peer support through a survey.

Methods: Following the Andersen and Newman Framework of Health Services Utilization and using REDCap (Research Electronic Data Capture), we designed and administered a survey to investigate the perceptions and utilization of online peer support among informal dementia caregivers. Specifically, we collected types of information that influence whether an informal dementia caregiver accesses online peer support: predisposing factors, which refer to the sociocultural characteristics of caregivers, relationships between caregivers and people living with dementia, and belief in the value of online peer support; enabling factors, which refer to the logistic aspects of accessing online peer support (eg, eHealth literacy and access to high-speed internet); and need factors, which are the most immediate causes of seeking online peer support. We also collected data on caregivers’ experiences with accessing online communities. We distributed the survey link on November 14, 2022, within two online locations: the Alzheimer’s Association website (as an advertisement) and ALZConnected (an online community organized by the Alzheimer’s Association). We collected all responses on February 23, 2023, and conducted a regression analysis to identify factors that were associated with accessing online peer support.

Results: We collected responses from 172 dementia caregivers. Of these participants, 140 (81.4%) completed the entire survey. These caregivers were aged 19 to 87 (mean 54, SD 13.5) years, and a majority were female (123/140, 87.9%) and White (126/140, 90%). Our findings show that the behavior of accessing any online community was significantly associated with participants’ belief in the value of online peer support ($P=.006$). Moreover, of the 40 non–online community caregivers, 33 (83%) had a belief score above 24—the score that was assigned when a neutral option was selected for each belief question. The most common reasons for not accessing any online community were having no time to do so (14/140, 10%) and having insufficient online information–searching skills (9/140, 6.4%).

Conclusions: Our findings suggest that online peer support is valuable, but practical strategies are needed to assist informal dementia caregivers who have limited time or online information–searching skills.

(JMIR Aging 2024;7:e55169) doi:10.2196/55169

KEYWORDS

informal dementia caregiver; online health community; social support; survey; online peer support; caregiving challenges
Introduction

Alzheimer disease—the most common cause of dementia—is a brain disorder that affects the thinking, comprehension, and learning capacity of more than 6 million Americans and is the seventh leading cause of death in the United States [1]. An estimated 80% of people living with Alzheimer disease or related dementia are cared for by unpaid informal caregivers (eg, family members, friends, or other unpaid caregivers) [2]. In 2021, over 11 million informal dementia caregivers provided 16 billion hours of care to people living with dementia [1]. Although this care was valued at nearly US $271.6 billion, it imposed substantial physical, financial, and mental burdens on these informal caregivers [3]. Additionally, 30% of informal dementia caregivers are aged ≥65 years [1]. They are likely to experience reduced social engagement due to caring for people living with dementia, which increases their risk of developing Alzheimer disease or some other dementia [4-6] and their risk of early death [7,8]. To ensure sufficient support for both informal caregivers and care recipients, it is essential for society to develop effective support mechanisms for the needs of informal dementia caregivers [9].

Research on how to best support informal dementia caregivers has focused primarily on assistance from credentialed professionals [10-12]. This type of assistance can improve a caregiver’s emotional well-being and caring skills, but maintaining this assistance over time can be difficult to achieve on a large scale. This is due, in part, to an insufficiently sized workforce, limited financial support [13,14], the stigma of asking for help, and difficulties encountered when leaving individuals with dementia [15,16]. In addition, if they lack a shared experience, it may be difficult for health care professionals or other family members to respond to the specific needs of informal caregivers. The perception that “they simply do not understand” [17] can contribute to feelings of loneliness [18], which were found to be negatively associated with the health and well-being of these caregivers [19].

The integration of the internet into daily life has enabled many people, including informal caregivers, to discuss health-related topics on online social media platforms [20,21]. For example, ALZConnected [22], which is organized by the Alzheimer’s Association, is the largest online community for people living with Alzheimer disease or related dementia and their caregivers in North America. ALZConnected has accumulated tens of thousands of online community users to discuss a broad range of topics regarding dementia caregiving and disease management [23]. Through online communities, informal dementia caregivers seek support and are willing to share experiences and practical information that they believe will assist other caregivers [24]. A study that analyzed an Alzheimer caregiver group on Facebook found that online peer social support had decreased the caregivers’ burdens while increasing their emotional and informational well-being [3]. Similarly, a survey found that increased online activity among caregivers was associated with lower levels of depression and loneliness [25]. In addition, many online communities provide added benefits, such as anonymity, asynchronous participation, and connection to numerous caregivers without physical location and time constraints [26], which provide cost-effective and convenient ways for informal dementia caregivers to gain support and access resources.

Based on internet utilization, informal dementia caregivers can be categorized as (1) non-internet caregivers, who never use the internet; (2) non–online community caregivers, who use the internet but do not participate in online communities; and (3) online community caregivers, who both use the internet and participate in online communities. Current social media–based dementia caregiving research primarily focuses on online community caregivers [25-27]. Although improving the online experiences of these caregivers is important [28], understanding how non–online community caregivers perceive the value of online peer support is significant as well. This would inform the development of interventions for non–online community caregivers to use and benefit from online peer support, thereby helping to mitigate the potential digital divide and decreasing existing health disparities in accessing online peer support [29].

The primary objective of this research was to gain insight into the perceptions and utilization of online community support among informal dementia caregivers, specifically non–online community caregivers and online community caregivers. To do so, we designed and administered a survey based on the Andersen and Newman Framework of Health Services Utilization (ANFHSU) [30]. The ANFHSU is a classical model for identifying and describing the factors that may affect a person’s access to and utilization of health services. Within this framework, we can analyze the various factors gathered through a survey questionnaire to determine an informal dementia caregiver’s access to or utilization of online peer support. Our findings suggest that online peer support is valuable, but practical strategies are needed to assist caregivers with limited time or online information–searching skills. This investigation marks the initial step toward addressing a long-term research objective, aiming to comprehensively elucidate the intricate mechanisms underlying online social support in dementia caregiving.

Methods

Online Peer Support

We begin this section with the definition of online peer support. Although peer support groups can be organized in an online format (eg, via Zoom meetings), the online peer support in this paper refers to the communications among informal dementia caregivers who may not know each other in real life but connect in online communities, forums, or websites (eg, Twitter, Facebook, Reddit, or ALZConnected). More broadly, since reading online caregiving discussions from other caregivers can serve as a way to obtain information and resources or learn caregiving skills, we also treated reading online posts as an online peer support–seeking behavior.

Ethical Considerations

This study was reviewed and deemed as non-human subjects research by the Vanderbilt University Medical Center Institutional Review Board (IRB 221732). Each survey participant was asked to confirm that they were unpaid informal dementia caregivers and that they were voluntarily participating.

https://aging.jmir.org/2024/1/e55169

Yin et al JMIR AGING 2024 | vol. 7 | e55169 | p.537

 JMIR Aging 2024 | vol. 7 | e55169 | p.337

/page number not for citation purposes)
in this survey in an information sheet. The sheet also included the definitions of informal Alzheimer disease and related dementia caregivers and online communities; the research purpose; a notice that the data would be collected, processed, and used; the methods applied to protect participant privacy; and a contact number for the Institutional Review Board.

**Questionnaire Design**

Social support is valuable in improving personal health [31]. In this study, we designed the survey based on the ANFHSU, whereby an informal dementia caregiver’s access to or utilization of online peer support was considered to be a function of 3 characteristics (Figure 1).

**Figure 1.** The predisposing, enabling, and need factors in the ANFHSU. ANFHSU: Andersen and Newman Framework of Health Services Utilization.

_Predisposing factors_ were the sociocultural characteristics of individuals or other factors that existed before these individuals became dementia caregivers. Specifically, we included the following predisposing factors:

- **Sociodemographics:** these included age, biological sex, education level, race, ethnicity, occupational status, and income; we also inquired about the number of children being cared for (if applicable), acknowledging the significant challenges in life balance experienced by caregivers in the “sandwich generation” [23].
- **Relationship between the caregiver and a person living with dementia:** this was included because caregivers who have different relationships with a person living with dementia face different caregiving challenges and burdens.
Belief in the value of online peer support: we asked the following seven 7-point Likert scale (ranging from “Strongly disagree” to “Strongly agree”) questions: “Do you believe that reading online discussions from, or directly writing posts to discuss with, other caregivers whom you do not know in the real world will help a) find caregiving resources that you need, b) increase caregiving knowledge, c) increase understanding of the disease and patient, d) improve caregiving skills, e) increase confidence in caregiving, f) reduce caregiving stress or g) reduce the feeling of loneliness as a caregiver?”

Health of the caregiver: we collected data on this factor through the following 7-point Likert scale (ranging from “Extremely disagree” to “Extremely agree”) item: “I feel healthy and do not have any major diseases that affect my daily life.”

Enabling factors referred to the logistic aspects of accessing online peer support. The following enabling factors were included:

- Access to a high-speed network, which is an important contributor to the digital divide [32].
- eHealth literacy: this was measured through the eHealth Literacy Scale (eHEALS) [33] to evaluate if a caregiver could process online health-related information.

Need factors referred to the most immediate causes of seeking online peer support. The following need factors were included:

- Health of the person living with dementia: for simplicity, caregivers reported the dementia stage for people living with dementia as “early,” “middle,” or “late.”
- Care involvement: this was measured through two questions (“How long have you been taking care of the PLWD (caregiving duration)?” and “How frequently do you care for the PLWD per week (caregiving workload)?”), and this design was based on the facts that dementia caregiving is a long-term dynamic process and that the weekly caregiving workload would affect the need for online peer support.
- Caregiving challenges: data on these were obtained via an open-ended question that requested caregivers to indicate the most challenging issues that they have faced when caring for a person living with dementia.
- Offline support sources: we also asked caregivers to describe (in a textbox) where to obtain support in the real world to handle the aforementioned challenges.
- Satisfaction with any support received in offline environments (7-point Likert scale).
- Caregiving stress: improving psychological well-being is a key focus of many interventions. The intuition is that higher levels of stress in caregivers often drive them to seek support. We measured this factor by using the Zarit Burden Interview (12-item) scale [34].
- Caregiver competence: this factor referred to one’s self-evaluation of their capacity to care for a person living with dementia and was measured by using the CARERS (Coaching, Advocacy, Respite, Education, Relationship, and Simulation) Interview (4-item) scale [35].

The rationale of the design of the survey was that if an informal dementia caregiver finds offline support to be insufficient in solving their caregiving challenges, they may turn to online environments for peer support. In addition to the ANFHSU factors, we surveyed each participant’s experiences with using online peer support in the past 3 months (ie, their behavior). For the survey participants who were already online community users, we asked the following questions: (1) “Which online platforms have you visited in the past three months?” (2) “How frequently did you visit those online communities?” (3) “Will you intend to revisit these online communities in the next three months (intention)?” (4) “If the answer to question 3) is yes, what are your motivations for revisiting these online communities?” For participants who were not online community users, we asked them an open-ended question regarding why they did not seek online peer support in the past 3 months.

Implementation and Dissemination

We implemented the survey questionnaire in REDCap (Research Electronic Data Capture; Vanderbilt University) [36]—a secure web application for building and managing online surveys. We distributed the survey link within 2 online locations. The first was ALZConnected, where we posted the survey link at the top of the thread board in the community’s two major caregiver forums: (1) Caregivers Forum and (2) Spouse or Partner Caregiver Forum. The second was the Alzheimer’s Association website [37], where an advertisement for this survey was created to allow any person who visited the website to have a chance to access the survey link. Given that this website is a popular source of information for patients with dementia and dementia caregivers, we expected to obtain survey responses from caregivers who were not part of online communities.

Analysis

We performed three types of analyses. First, we summarized the answers to the multiple-choice questions by illustrating the respondent distribution for each choice. This was done to paint a broad picture of who the responding caregivers were and their perceptions and utilization of online peer support.

Second, we summarized the answers to the open-ended questions through manual annotation. This was not a trivial task because the responses were written as free text, and there were no predefined categories before the annotation. To address these issues, two authors (CN and QS) read the responses and annotated the categories for each open-ended question independently. Next, both annotators compared and discussed their summarized categories to create the final categories, including category names and definitions. Both annotators then independently modified their annotations with the agreed upon categories. For each response, we adopted a conservative approach and only reported the interactions of the categories summarized by the two annotators.

Finally, we fitted a logistic regression model (R v4.2.2; R Foundation for Statistical Computing) to analyze how the proposed ANFHSU factors were associated with using online peer support. Due to the small sample size, we converted some categorical variables into binary ones. Specifically, we encoded gender as “female/non-female,” race as “white/non-white,” education level as “4-year college degree or above/below 4-year college degree,” access to high-speed network as “yes/no,”...
relationship as “spouse/non-spouse,” and all the other
categorical variables (eg, annual income, dementia stage,
caregiving duration, and caregiving workload) as ordinal
variables (eg, numerical values with equal distance). A 2-sided
P value of <.05 was considered to be statistically significant.

Results

Survey Overview
We distributed the survey link on November 14, 2022, leaving
it open for over 3 months until February 23, 2023. During this
time period, we collected responses from 172 dementia
caregivers. Of these participants, 140 (81.4%) completed the
entire survey.

Characteristics of Caregivers and People Living With
Dementia

Table 1 summarizes the sociodemographic and caregiving
characteristics of the 140 caregivers who completed the survey.
These caregivers were aged 19 to 87 (mean 54, SD 13.5) years,
and a majority were female (123/140, 87.9%) and White
(126/140, 90%). Over 60% (89/140, 63.6%) of these caregivers
had a 4-year college degree or higher, whereas 36.4% (51/140)
had a below–4-year college education level. A majority of the
caregivers were the adult children of the people living with
dementia (71/140, 50.7%), followed by spouses or partners
(41/140, 29.3%) and other relatives (19/140, 13.6%). Only
40.7% (57/140) of these caregivers were employed full-time,
and 28.6% (40/140) took care of the person living with dementia
and at least one child simultaneously. Further, 17.9% (25/140)
and 32.1% (45/140) of the respondents reported taking care of
the person living with dementia for <1 year and >4 years,
respectively, and 66.4% (93/140) provided daily care. The
caregivers’ annual income was approximately uniformly
distributed across incomes of less than US $25,000 to incomes
ranging between US $100,000 and US $149,999. Only 7.9%
(11/140) of these caregivers earned more than US $149,999.

Table 2 presents the characteristics of the people living with
dementia reported by 140 survey participants. The ages of the
people living with dementia ranged from 46 to 97 (mean 76,
SD 9.5) years, and 65% (91/140) were female. Over 60%
(86/140, 61.4%) of the people living with dementia were at the
middle stage of dementia, 21.4% (30/140) were at the late stage,
and 17.1% (24/140) were at the early stage.
<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Caregivers (N=140)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (y), mean (SD; range)</strong></td>
<td>54 (13.5; 19-87)</td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>123 (87.9)</td>
</tr>
<tr>
<td>Male</td>
<td>16 (11.4)</td>
</tr>
<tr>
<td>Undifferentiated</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td><strong>Race, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>126 (90)</td>
</tr>
<tr>
<td>Asian</td>
<td>8 (5.7)</td>
</tr>
<tr>
<td>Black or African American</td>
<td>5 (3.6)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td><strong>Ethnicity, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Not Hispanic or Latino</td>
<td>130 (92.9)</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>10 (7.1)</td>
</tr>
<tr>
<td><strong>Education level, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Above a 4-y college degree</td>
<td>57 (40.7)</td>
</tr>
<tr>
<td>4-y college graduate</td>
<td>32 (22.9)</td>
</tr>
<tr>
<td>Some college or 2-y degree</td>
<td>40 (28.6)</td>
</tr>
<tr>
<td>High school or General Educational Develop-</td>
<td>11 (7.9)</td>
</tr>
<tr>
<td>m</td>
<td></td>
</tr>
<tr>
<td><strong>Employment status, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Full time</td>
<td>57 (40.7)</td>
</tr>
<tr>
<td>Retired</td>
<td>40 (28.6)</td>
</tr>
<tr>
<td>Part time</td>
<td>22 (15.7)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>21 (15)</td>
</tr>
<tr>
<td><strong>Relationship, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Adult child</td>
<td>71 (50.7)</td>
</tr>
<tr>
<td>Spouse or partner</td>
<td>41 (29.3)</td>
</tr>
<tr>
<td>Other relative</td>
<td>19 (13.6)</td>
</tr>
<tr>
<td>Grandchild</td>
<td>6 (4.3)</td>
</tr>
<tr>
<td>Neighbor</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Friend</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Other</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td><strong>Number of children cared for, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>15 (10.7)</td>
</tr>
<tr>
<td>2</td>
<td>20 (14.3)</td>
</tr>
<tr>
<td>3</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td>&gt;3</td>
<td>3 (2.1)</td>
</tr>
<tr>
<td>Does not apply</td>
<td>100 (71.4)</td>
</tr>
<tr>
<td><strong>Caregiving duration, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>&lt;6 mo</td>
<td>12 (8.6)</td>
</tr>
<tr>
<td>6-12 mo</td>
<td>13 (9.3)</td>
</tr>
<tr>
<td>1-2 y</td>
<td>35 (25)</td>
</tr>
</tbody>
</table>
Table 1. Characteristics of the people living with dementia reported by survey participants.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>People living with dementia (N=140)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y), mean (SD; range)</td>
<td>76 (9.5; 46-97)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>91 (65)</td>
</tr>
<tr>
<td>Male</td>
<td>49 (35)</td>
</tr>
<tr>
<td>Dementia stage, n (%)</td>
<td></td>
</tr>
<tr>
<td>Early stage</td>
<td>24 (17.1)</td>
</tr>
<tr>
<td>Middle stage</td>
<td>86 (61.4)</td>
</tr>
<tr>
<td>Late stage</td>
<td>30 (21.4)</td>
</tr>
</tbody>
</table>

Caregiving Challenges and Support Sources

Table 3 summarizes reported caregiving challenges and the sources where caregivers sought support. The main caregiving challenges included dealing with the memory issues of a person living with dementia (40/140, 28.6%), supporting a person living with dementia in their daily life (such as showering and transportation; 30/140, 21.4%), and maintaining a balanced life (24/140, 17.1%). It should be noted that life balancing included the ability to (1) balance taking care of one’s children and a person living with dementia, (2) balance work and caregiving, and (3) balance one’s social life and caregiving. Dealing with the emotional fluctuations of the person living with dementia (21/140, 15%) and dealing with financial issues (14/140, 10%) were also major caregiving challenges. Some caregivers expressed concerns about the dementia treatment (11/140, 7.9%) for people living with dementia and their own mental health issues (10/140, 7.1%).
### Table. Summary of caregiving challenges and specific support sources.\(^a\)

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Caregivers (N=140), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legal issue</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Physical health</td>
<td>1 (0.7)</td>
</tr>
<tr>
<td>Other health issue</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td>Mental health (caregiver)</td>
<td>10 (7.1)</td>
</tr>
<tr>
<td>Treatment concern</td>
<td>11 (7.9)</td>
</tr>
<tr>
<td>Family conflict</td>
<td>13 (9.3)</td>
</tr>
<tr>
<td>Financial issue</td>
<td>14 (10)</td>
</tr>
<tr>
<td>Emotion change</td>
<td>21 (15)</td>
</tr>
<tr>
<td>Life balancing</td>
<td>24 (17.1)</td>
</tr>
<tr>
<td>Daily caregiving</td>
<td>30 (21.4)</td>
</tr>
<tr>
<td>Memory issue</td>
<td>40 (28.6)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Support source</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Self-learning (book)</td>
<td>4 (2.9)</td>
</tr>
<tr>
<td>Dementia caregivers (offline)</td>
<td>7 (5)</td>
</tr>
<tr>
<td>Self-learning (online search)</td>
<td>8 (5.7)</td>
</tr>
<tr>
<td>Local support group</td>
<td>23 (16.4)</td>
</tr>
<tr>
<td>Friend</td>
<td>35 (25)</td>
</tr>
<tr>
<td>Dementia caregivers (online community)</td>
<td>35 (25)</td>
</tr>
<tr>
<td>Health care provider</td>
<td>38 (27.1)</td>
</tr>
<tr>
<td>Family member</td>
<td>66 (47.1)</td>
</tr>
</tbody>
</table>

\(^a\)Caregivers could report multiple challenges and support sources, which is why the total of the percentages for each section exceed 100%.

### Satisfaction With Support and Network Access

Table 4 depicts the distribution of caregivers by satisfaction with the support they sought and by their access to a high-speed network. Of the 140 caregivers, 69 (49.3%) reported an “agree” or above response (which included “Slightly agree,” “Quite agree,” and “Extremely agree”) for their satisfaction with the sought support. Among the 56 (40%) caregivers who reported negative experiences, 12 (8.6%) reported “Extremely disagree,” indicating the challenging situations faced by these caregivers. Only 6 caregivers reported that they could not access a high-speed network.
Table. Distribution of caregivers by satisfaction with support that was sought and by access to a high-speed network.

<table>
<thead>
<tr>
<th>Items and responses</th>
<th>Caregivers, n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Satisfaction with the sought support</strong></td>
<td></td>
</tr>
<tr>
<td>Extremely disagree</td>
<td>12</td>
</tr>
<tr>
<td>Quite disagree</td>
<td>29</td>
</tr>
<tr>
<td>Slightly disagree</td>
<td>15</td>
</tr>
<tr>
<td>Neither</td>
<td>15</td>
</tr>
<tr>
<td>Slightly agree</td>
<td>37</td>
</tr>
<tr>
<td>Quite agree</td>
<td>31</td>
</tr>
<tr>
<td>Extremely agree</td>
<td>1</td>
</tr>
<tr>
<td><strong>Access to a high-speed network</strong></td>
<td></td>
</tr>
<tr>
<td>Extremely disagree</td>
<td>1</td>
</tr>
<tr>
<td>Quite disagree</td>
<td>2</td>
</tr>
<tr>
<td>Slightly disagree</td>
<td>3</td>
</tr>
<tr>
<td>Slightly agree</td>
<td>11</td>
</tr>
<tr>
<td>Quite agree</td>
<td>38</td>
</tr>
<tr>
<td>Extremely agree</td>
<td>85</td>
</tr>
</tbody>
</table>

Measures of Caregiving Stress and Competence, eHEALS, and Belief

Figure 2 illustrates the distribution of the scores associated with caregiving stress, caregiving competence, eHealth literacy, and a caregiver’s belief in the value of peer support from an online environment. The Cronbach α—a measure of the internal consistency of a questionnaire or survey (the higher, the better)—of these measures was 0.86 (95% CI 0.82-0.89), 0.83 (95% CI 0.78-0.87), 0.87 (95% CI 0.84-0.90), and 0.87 (95% CI 0.83-0.89), respectively. These indicated very good internal consistency. Specifically, of the 140 caregivers, 121 (86.4%) had a stress score above 30 (the score when “neutral” was selected for all the questions; the same cutoffs were used for each measure), 67 (47.9%) had a competence score above 12, 138 (98.6%) had an eHealth literacy score above 20, and 125 (89.3%) had a belief score above 24. These results suggested that most of the caregivers were in stressful caregiving experiences, had high eHealth literacy, and believed in the value of online peer support. However, over half (73/140, 52.1%) of these caregivers were not confident in their caregiving skills.

Figure 3 provides a detailed illustration (in the form of heat maps) of the responses for each belief category and the correlations in the responses. It shows that obtaining the needed resources, increasing one’s understanding of the patient and the disease, and improving caregiving skills and confidence were the top “believed” values of online peer support. In contrast, reducing caregiving stress and the loneliness of being a caregiver was slightly more challenging. However, both categories received more “likely” responses than “unlikely” responses, with 57.1% (80/140) and 65% (91/140) of caregivers selecting “likely” responses, respectively. There were also several interesting observations made. First, reducing loneliness as a caregiver was highly correlated with reducing stress. Second, improving caregiving skills was highly correlated with increasing caregiving confidence. Third, reducing stress was correlated with improving caregiving skills and increasing caregiving confidence. Although correlation does not necessarily imply causation, these results suggested that improving caregiving skills and reducing loneliness may help to reduce caregiving stress.
Figure 2. Distribution of scores for the four measures. The red vertical lines in the “Caregiving stress,” “eHealth literacy,” and “Belief” in online peer support graphs correspond to the scores when “neutral” or “undecided” was selected for all the questions in each measure. The red vertical line in the “Caregiving competence” graph corresponds to the score when the “fairly” option was selected for all the questions.
Utilization of Online Peer Support

Of the 140 survey participants, 40 (28.6%) reported never using any online community to either read online caregiving discussions or discuss caregiving issues with other online peers in the past 3 months. The regression analysis showed that only the belief score was statistically significantly associated with the utilization of online peer support (β: 0.11, SD 0.041; P=.006). All the other factors, including sociodemographics, care duration and workload, dementia stage, access to a high-speed network, eHealth literacy, stress score, and competence score, did not have statistically significant effects (all P values were >.10). Additionally, of the 40 non–online community caregivers, 33 (83%) had a belief score above 24—the score when “neutral” was selected for all the questions. The reasons for not using any online community to access online peer support included having no time for online activities due to the intensive caregiving workload (14/140, 10%), having a lack of searching skills (9/140, 6.4%), having belief in unreliable online information (6/140, 4.3%), and having security and privacy concerns (2/140, 1.4%). Further, 4 (2.9%) caregivers reported not wanting to spend time online after working on a computer during the daytime. Interestingly, 99 of the 100 online community caregivers reported an intention to revisit online communities in the following 3 months. The only caregiver...
who did not do so said that too many sad stories in online communities made her worry about her father’s future.

**Discussion**

**Principal Results**

When comparing with *Alzheimer’s Disease Facts and Figures* (AFF) [1], we observed that our survey recruited a similar proportion of caregivers who were Hispanic caregivers (our survey: 10/140, 7.1%; AFF: 8%) and Asian caregivers (our survey: 8/140, 5.7%; AFF: 5%). However, the proportion of Black or African American caregivers we recruited was far below that reported in AFF (our survey: 5/140, 3.6%; AFF: 10%), which suggests that Asian and Hispanic caregivers are more likely to participate in research studies when compared to Black caregivers [38]. To gain greater insight into the situation, we reviewed the responses of the five caregivers who reported their race as Black. We found that two of the caregivers accessed online communities in the past 3 months, with one (age: 49 y; education level: college degree; stress score: 51) reporting this at least once per week and the other (age: 50 y; education level: more than a 4-y college degree; stress score: 19) reporting doing this when she had a caregiving question that needed an answer. Among the three remaining caregivers who reported their race as Black and did not access any online community in the past 3 months, one (age: 60 y; education level: 4-y college degree; stress score: 55) said, “I have used books and I think some information on the internet can be misinterpreted”; another one (age: 29 y; education level: some college or 2-y degree; stress score: 55) said, “I have recently started looking for online communities”; and the third (age: 30 y; education level: some college or 2-y degree; stress score: 52) said that they were “unaware.” However, all their belief scores were above 24 and ranged from 25 to 42, which suggests that these three caregivers believed in the value of online peer support, but they may not have known how to search online communities for peer support. Notably, the caregiver who relied on books for information seeking exhibited the lowest belief score (ie, a score of 25), indicating the survey results’ reliability.

Our study shows that the most prevalent challenges faced by these informal caregivers were (1) handling the memory issues of a person living with dementia, (2) daily caregiving, (3) life balance, and (4) emotional fluctuations. Caregivers facing such challenges can, at times, receive guidance from their peers who have shared the same firsthand experiences. This suggests that online communities, such as ALZConnected, that provide an online platform for caregivers to discuss their challenges and experiences can be a valuable resource for those seeking informational or emotional peer support. This implication is further supported by our observation that a large proportion of the surveyed caregivers indicated a belief in the value of online peer support. Still, given that some caregivers did not access online environments for peer support, it is evident that online environments alone are not a comprehensive solution and that caregivers need other types of support. Moreover, there are various caregiving challenges that are more likely to be addressed by professionals rather than peers, such as clinical questions about treatment, legal issues, and financial concerns.

Given that 50.7% (71/140) of the survey participants were not satisfied with their received support, it is essential to diversify the support source portfolio to assist a broad range of caregivers. As such, the use of online environments can serve as one of a collection of strategies that, in concert with one another, provide a support structure for informal caregivers.

Another main result of this study is that whether an informal dementia caregiver accessed online health communities in the past 3 months depended upon their belief in the value of peer support obtained from online health communities but not upon their sociodemographics or any other ANFHSU factors. For example, one caregiver said, “I don’t find comfort from strangers on the internet, I would love to, and I am willing to go to an in-person meeting of a support group, but I don’t have anyone to watch my mother so that I could attend,” and they exhibited a belief score of 14, which is logical based on their reason for not accessing any online communities. This further highlights the dilemma that there is limited time to attend local support groups because of intensive caregiving [15]. Another caregiver, who had a belief score of 14, indicated that their health insurance could only cover 40 days of at-home nursing support in 1 year, which made their family feel so “helpless and alone.” Despite various reasons for not using online health communities in the past 3 months, 89.3% (125/140) of respondents exhibited a belief score above 24. This suggests that online peer support was valuable to them, but an effective strategy for bridging their needs and the desired online peer support is needed. Finally, the correlation of the survey responses to the six value aspects implied that an intervention designed around peer learning can effectively enhance a caregiver’s caregiving competence and reduce feelings of loneliness.

**Limitations**

There are, however, several limitations to this study that can serve as a basis for future research. First, since we distributed our survey link in the ALZConnected online community and the Alzheimer’s Association’s website, our results may be biased toward online community users. A less biased approach may be designed to collect data that reflect the perceptions and utilization of online peer support in the dementia caregiver population. Making online peer support beneficial to noninternet caregivers is equally essential, but determining how to address the internet access issue is a priority and is beyond the scope of this study. Second, only a small percentage of the participants were caregivers who reported their race as Black. Increasing participation in this group would increase the understanding of their perceptions and utilization of online peer support in the dementia caregiver population. Making online peer support beneficial to noninternet caregivers is equally essential, but determining how to address the internet access issue is a priority and is beyond the scope of this study. Second, only a small percentage of the participants were caregivers who reported their race as Black. Increasing participation in this group would increase the understanding of their perceptions and utilization of online peer support. Third, the analysis relied on 140 completed responses, indicating limited statistical power. Although statistically significant findings emerged, larger sample sizes are necessary for a more precise examination of this research. Fourth, in an open-ended question, we inquired about the types of offline support caregivers received to address their caregiving challenges. Some caregivers mentioned online peer support in their responses, leading to misalignment between their answers and the original research design, particularly regarding satisfaction with offline support. Future investigations should clarify this question or convert this question to a multiple-choice version. Finally, it is essential to study how to help caregivers without time or...
sufficient online information—searching skills screen the online caregiving discussions they need.

Conclusions
This study reported on an online survey about the perceptions and utilization of online peer support among informal dementia caregivers. Belief in the value of online peer support was statistically significantly associated with accessing online communities ($P=.006$). Moreover, there were a number of caregivers who were not using online peer support but held belief in the value of such support. As such, there is clearly an opportunity to build tools that help caregivers who are existing online community users and caregivers who intend to seek online information find reliable, matched online peer support.

Acknowledgments
ZY and L Stratton designed the survey questionnaire. L Song provided the theoretical support for the survey design. L Song and PC revised the survey questionnaire. QS collected and cleaned the survey responses. ZY drafted the first version of the manuscript. All the other authors edited and approved the manuscript.

Conflicts of Interest
None declared.

References


36. REDCap. Research Electronic Data Capture. URL: https://www.project-redcap.org [accessed 2024-05-06]


Abbreviations

AFF: 2022 Alzheimer’s Disease Facts and Figures
ANFHSU: Andersen and Newman Framework of Health Services Utilization
CARERS: Coaching, Advocacy, Respite, Education, Relationship, and Simulation
eHEALS: eHealth Literacy Scale
REDCap: Research Electronic Data Capture
Abstract

**Background:** Social media platforms have been effective in raising awareness of the underassessment and undertreatment of pain in dementia.

**Objective:** After a successful pilot campaign, we aimed to scale our pain-in-dementia knowledge mobilization pilot initiative (ie, #SeePainMoreClearly) to several social media platforms with the aid of a digital media partner. The goal of the initiative was to increase awareness of the challenges in the assessment and management of pain among people with dementia. A variety of metrics were implemented to evaluate the effort. Through this work, we endeavored to highlight key differences between our pilot initiative (which was a grassroots initiative), focusing largely on Twitter and YouTube, and the current science-media partnership. We also aimed to generate recommendations suitable for other social media campaigns related to health or aging.

**Methods:** Evidence-based information about pain in dementia was summarized into engaging content (eg, videos) tailored to the needs of various knowledge users (eg, health professionals, families, and policy makers). We disseminated information using Facebook (Meta Platforms), Twitter (X Corp), YouTube (Alphabet Inc), Instagram (Meta Platforms), and LinkedIn (LinkedIn Corp) and measured the success of the initiative over a 12-month period (2020 to 2021). The evaluation methods focused on web analytics and questionnaires related to social media content. Knowledge users’ web responses about the initiative and semistructured interviews were analyzed using thematic analysis.

**Results:** During the course of the campaign, >700 posts were shared across all platforms. Web analytics showed that we drew >60,000 users from 82 countries to our resource website. Of the social media platforms used, Facebook was the most effective in reaching knowledge users (ie, over 1,300,000 users). Questionnaire responses from users were favorable; interview responses indicated that the information shared throughout the initiative increased awareness of the problem of pain in dementia and influenced respondent behavior.

**Conclusions:** In this investigation, we demonstrated success in directing knowledge users to a resource website with practical information that health professionals could use in patient care along with pain assessment and management information for
caregivers and people living with dementia. The evaluation metrics suggested no considerable differences between our pilot campaign and broader initiative when accounting for the length of time of each initiative. The limitations of large-scale health campaigns were noted, and recommendations were outlined for other researchers aiming to leverage social media as a knowledge mobilization tool.

*JMIR Aging 2024;7:e53025*  doi:10.2196/53025

**KEYWORDS**

knowledge translation; Twitter; older adults; Facebook; knowledge mobilization

**Introduction**

**Background**

Social media platforms play an important role in academic dissemination and have untapped potential as knowledge mobilization (KM) tools that can allow researchers to interact directly with the public worldwide [1]. KM encompasses activities involved in the synthesis and dissemination of research information; this process includes knowledge synthesis, dissemination, transfer, exchange, and cocreation by researchers and knowledge users [2]. Knowledge users, such as older adults and their families, are increasingly turning to web-based platforms (eg, Twitter [subsequently rebranded as X] and Facebook) to share and obtain information [3,4]. Health-related initiatives on social media networks have successfully raised awareness about a host of topics (eg, mental health, breastfeeding, and cancer) [5]. For example, a nationwide HIV prevention campaign (#PrEP4Love) garnered >40 million views across various social media platforms [6].

Researchers have supplemented social media KM initiatives with other web-based KM methods (eg, web-based repositories, educational videos, and community groups) [7,8]. Despite the innovation of this approach, the sustainability of social media health campaigns is contingent upon different factors (eg, continued content generation and ongoing funding) [9]. As such, there is a growing need for partnerships with digital media experts to aid in health KM. A notable example of this is the work by Chambers [10], who mobilized knowledge about pediatric pain within the context of a science-media partnership through her #ItDoesntHavetoHurt campaign. The initiative demonstrated worldwide impact and great success in producing and disseminating digital evidence-based content tailored to the needs of parents and health professionals, obtaining 1.3 million Twitter impressions, 5.5 million hashtag mentions, and >250,000 views of her YouTube video [10]. Impact was also demonstrated through indices of adoption and changed practices [10].

Although researchers have attempted to track the reach and impact of social media for KM through web metrics and the use of social listening software [11,12], very little research has been conducted using social media for KM in older adults with dementia. To address limitations in the reach of traditional KM campaigns and a gap in the literature, we launched and evaluated the pilot #SeePainMoreClearly (phase I) social media KM campaign with messages that reached >2,576,853 unique individuals on Twitter [9]. The goal of the pilot initiative was to increase awareness of the challenges in the assessment and management of pain among people with moderate to severe dementia with limited ability to communicate [13,14]. Moreover, we aimed to disseminate information on evidence-based approaches to effectively assess and manage pain in this population (eg, [15-18]). To maximize uptake, we created a web-based repository of pain assessment and management information [19]. We also prepared an engaging 2-minute YouTube video with evidence-based information about pain in dementia [20].

The script for the video was created and modified with input from health professionals, caregiver partners, researchers, and knowledge user organizations. The pilot campaign was evaluated by tracking social media and web metrics and by conducting qualitative analyses of social media posts in response to the #SeePainMoreClearly campaign over a 5-month period. #SeePainMoreClearly demonstrated substantial reach with >5,000,000 hashtag impressions on Twitter. The short video was viewed >50,000 times, and our web-based repository was visited by people in >55 countries. Moreover, the content analysis of social media posts (ie, tweets) from users who used the hashtag or responded to our messaging and content posted on Twitter were favorable. Many users expressed support for the initiative and increased advocacy for improved pain care for people living with dementia.

Our findings showed the effectiveness of web-based KM methods in reaching very broad international audiences quickly. Perhaps most importantly, without counting the posts that were produced by members of our team, the initiative doubled the number of posts made on the topic of pain dementia on Twitter during the campaign period as compared with a control period of the previous year [9]. It is important to note that the #SeePainMoreClearly pilot campaign took place in 2019, before the COVID-19 pandemic. The pilot #SeePainMoreClearly was a grassroots initiative in which a small number of researchers prepared engaging materials and disseminated them on their own with support from several knowledge user organizations that agreed to help disseminate the message. A lesson from the pilot #SeePainMoreClearly experience was that although KM grassroots campaigns can be very successful, lengthier social media campaigns (run solely by researchers) would not be sustainable, as involvement can be very time-consuming for researchers who generally must attend to multiple other obligations. A partner with digital media expertise and time to dedicate resources is necessary for a sustainable large-scale KM campaign. Hence, we sought and obtained funding to support such a partnership.
Objectives

The primary purpose of this study was to launch and evaluate a longer (12 months instead of 5 months) #SeePainMoreClearly campaign with expanded social media platform coverage (eg, Twitter, Instagram, Facebook, and LinkedIn as opposed to just Twitter) with the following additional objectives:

1. Track the reach of a larger social media KM campaign over a 12-month period on various platforms (eg, Facebook, Twitter, Instagram, and LinkedIn)
2. Evaluate the impact of the campaign on knowledge users’ (eg, patient, caregiver, health professional, and policy maker) knowledge
3. Evaluate the impact of a science-media partnership on our pilot campaign in meeting knowledge users’ needs
4. Outline recommendations to develop and scale social media KM initiatives.

We expected the results to demonstrate a large scope and breadth of our campaign, resulting in a large number of visits to our resource website [19]. Moreover, we predicted that our results would demonstrate the value of a science-media partnership as a strategy for improving KM in the area of pain in dementia. We contrast the experiences of the digital media partnership campaign with the pilot #SeePainMoreClearly campaign. The findings from this project can inform subsequent evaluations of social media as a KM tool.

Methods

Ethical Considerations

The evaluation process involving contact with human participants (eg, questionnaires and semistructured interviews) was approved by the University of Regina Research Ethics Board (#REB 2020-036).

Identification and Engagement of Digital Media Partner

Given the nature of our funding, we were restricted to identify a digital media partner within our home province. As such, a web search was conducted to identify potential locally based digital media partners that listed social media promotion as part of their services. We then communicated with 2 of the firms that, based on their websites, seemed highly experienced with social media marketing and described the nature of our work to confirm that they understood the needs of our project and that they were interested in partnering with us for the campaign. Our institution required us to subject the project to an open competitive bidding process before awarding the contract. This process allows qualified bidders to submit proposals and budgets. The 2 firms with which we had communicated were invited to submit their proposals. On the basis of this process, the best qualified bidder who offered competitive pricing was selected from a total of 3 bidders.

Development of Content and the Campaign

We worked with the selected digital media partner to develop the campaign content and KM strategy as well as to collect social media analytics in response to our campaign. An iterative process was implemented with team collaborators (eg, caregiver partners, researchers with knowledge of the content area, media experts, and health care professionals) with the aim of developing lay summaries of evidence-based information about pain in dementia. Specifically, a series of web-based meetings with members of the team were coordinated to identify pertinent information about pain assessment and management among older adults with dementia, which served as the basis of the campaign’s content. Accordingly, the general topics agreed upon by the team were as follows: (1) the feasibility of regular pain assessments, (2) available guidelines for pain assessment, (3) resources for informal caregivers, (4) effects of psychotropic medication in long-term care (LTC), (5) validated pain assessment tools, (6) implementing regular pain care in LTC, (7) gaps in education for health professionals and ways to address this challenge, (8) effects of untreated pain in dementia, (9) pain care during the COVID-19 pandemic, (10) cost of untreated pain for the health care system, (11) benefits of regular pain assessments, and (12) effective ways to assess pain in dementia. Once the general topics were identified, information was gathered that corresponded to established guidelines (eg, [17,18,21,22]) or was supported by published research in leading peer-reviewed journals. Specific evidence-based information was selected by our team, which included content experts and caregivers with lived experience. Next, the team developed lay summaries (eg, 200-300 words) of the evidence present in the literature covering each aforementioned topic (eg, [17,18,21,22]). Several rounds of refinement were conducted by team members. Following this, the team developed 8 cross-cutting messages and key points (Multimedia Appendix 1) divided by the target group (eg, health professionals, policy makers, and public or families). These messages were then relayed to the digital media team as the basis for the messaging developed for the campaign. The digital marketing team developed a series of images and posts for each target group. Moreover, a short animated (ie, 2-minute long) video was created for each target group (eg, family, health care professionals, researchers, and policy makers). The animated videos were posted on the web page and on social media platforms. Examples of the content and posts are shown in Figure 1. Blog posts written by team members and people with lived experiences related to the topics outlined in this section were shared throughout the initiative. A total of 42 blog posts [23] were created and posted on the web page of this project.
**Dissemination Strategy**

Unlike the pilot campaign [9], which focused on Twitter and YouTube, information was disseminated on 5 social media platforms used for the campaign: Facebook (Meta Platforms), Twitter (X Corp), YouTube (Alphabet Inc), Instagram (Meta Platforms), and LinkedIn (LinkedIn Corp).

A total of 756 posts were shared across Facebook, Twitter, Instagram, and LinkedIn, and 5 videos were posted on YouTube over a 12-month period (ie, October 1, 2020, to September 30, 2021). A key goal of the dissemination strategy was to direct knowledge users to the web-based repository to share further information and resources about pain in dementia [19]. We designed the content (eg, social media posts, images, and short animated videos) to be of interest to our target groups. Moreover, we developed social media posts to gain the attention of specific target groups (eg, by including an image with the statement *Resources for researchers* prominently displayed). As with the pilot initiative, the #SeePainMoreClearly hashtag was used to maximize the reach of the content and information. We leveraged events (eg, the International Day of Older Persons in October, World Alzheimer Awareness Month in September, and Alzheimer Awareness Month in January) to develop tailored content. The total amount of money spent on post promotion and advertisements was CAD $29,793.86 (US $22,133.06); the vast majority of this amount was spent on Facebook and Instagram, CAD $26,384 (US $19,599.96), Twitter, CAD $3045 (US $2262.05), and LinkedIn, CAD $364.86 (US $271.04).

**Evaluation of the Campaign**

**Engagement and Reach of the Initiative**

**Web Analytics**

A longitudinal data collection strategy was used to obtain web and social media metrics over a 12-month campaign period (ie, October 1, 2020, to September 30, 2021). For the website, the number of content viewers and pages views were tracked using Google Analytics [24]. For our analysis, we reported a descriptive count (eg, count of total engagement, reach, and impressions) of all web and social media analytics. Google Analytics was used to collect data from the website. To clarify how Google Analytics works, each time a user visits a web page, a tracking code by Google Analytics collects information on how that user interacted with the page [25]. Google Analytics then aggregates and organizes this information in the Google Analytics portal as a report detailing different metrics (eg, the count of web page visits and the number of pages viewed). Similarly, the social media metrics were aggregated by each social media platform. This information was accessed through...
the social media platform’s website (eg, Facebook Analytics). Web and social media analytics were exported and reported.

Social Media Analytics

The reach, scope, and engagement of the initiative and web-based repository were monitored through social media metrics (eg, web analytics, hashtag analytics, and social media analytics). Similar metrics have been used in previous research to quantify the impact of social media initiatives [10,12]. The number of impressions (ie, the number of times users have seen the web-based content), reach (ie, the number of unique users who have seen the web-based content), and engagement (ie, the number of comments, retweets, "likes," and shares) were extracted using the analytics provided by each platform. Not all analytics were available for each social media platform (eg, reach was not available for LinkedIn). Analytics were summarized based on three sources of information: (1) analytics of only paid posts provided by the social media platform, (2) analytics of the total number of posts (ie, paid and unpaid) provided by the social media platform, and (3) third-party social media analytics manager (eg, CloudCampaign) provided by the digital marketing team. Moreover, our digital media partner could only obtain data for specific periods (eg, year-long data could not always be obtained) for some platforms owing to limitations in the sources of information. Nevertheless, attempts were made to encompass the entire evaluation period. Depending on the data source, data were provided from (1) October 1, 2020, to September 3, 2021; (2) October 1, 2020, to November 2, 2020; or (3) November 3, 2020, to September 30, 2021. Thus, the reported numbers varied as a function of the period. The Keyhole social media monitoring (also known as social listening) software [26] was used to collect the metrics and analytics for the #SeePainMoreClearly hashtag on Twitter.

Knowledge Users’ Responses to the Initiative

Semistructured Interviews With Knowledge Users

Individuals who viewed and engaged with the content (eg, responded to social media posts or added their own commentary when reposting content shared throughout the initiative) on social media were invited to participate (ie, research personnel commented back to users who responded to social media posts and invited them to participate in interview) in semistructured interviews with a researcher to gain an understanding of their perceptions about the initiative. Moreover, invitations (ie, through social media posts or recruitment posters) to participate in the interview were circulated across social media platforms. All participants recruited through social media posts were asked to contact the research team to indicate their interest. The interviews were conducted over Zoom (Zoom Video Communications Inc) audio-only and covered the following topics: perceptions about the initiative, quality of information and messages disseminated, and impact of the initiative on knowledge and behavior (Multimedia Appendix 2). Zoom’s automated transcription feature was used to facilitate the transcription of the interviews. The transcription was then reviewed with the interview audio recording to ensure accuracy. NVivo (version 12; Lumivero) software [27] was used to facilitate the coding reliability thematic analysis [28]. Two independent researchers initially organized a subset of data into clusters based on commonalities and recurring ideas to develop a coding book [28]. Consistent with a coding reliability thematic analysis approach, themes were developed inductively and aligned with the data collection questions of the study [28]. Themes that emerge from this analysis can be viewed as a summary of the participant responses. Codes were developed after data familiarization, while keeping the data collection questions in mind. Two coders individually coded a subset (3/13, 20%) of the data to develop a coding book. They grouped each line of text into clusters and then into higher-level codes (ie, themes). The coders then met to discuss and finalize the coding book. Discrepancies between the coding researchers and decisions to merge or split categories were discussed throughout the process until a consensus was reached. The developed coding book was applied to the remaining data to identify the prevalent categories and obtain a frequency count of the identified themes. To ensure rigor in our analysis, we conducted consistency checks on the coded themes and categories (eg, discussing discrepancies between coders) and maintained ongoing communication among coders throughout the process to establish consensus. Moreover, a randomly selected subset of data was coded to assess intercoder agreement. Triangulation was considered through the different perspectives (eg, health professionals, caregivers, and people with dementia) obtained in this study and the use of a mixed methods approach (ie, use of interview and numeric questionnaire data). A primary coder coded the entire data set, and a secondary coder coded a subset of the data to establish intercoder agreement. The second coder coded a randomly selected 20% (3/13) of the participant data. NVivo calculates a κ coefficient to evaluate agreement as follows: total units of agreement between the 2 coders minus the expected frequency of the agreement occurring by chance, divided by the total units within the source minus the expected frequency of the agreement occurring by chance [29].

Social Media Comments About the Campaign

Social media analytics do not provide information about the content of the comments in response to our initiative (ie, whether users liked or disliked the content). As such, we analyzed the responses to the initiative to better understand users’ perception of our campaign. To assess the perception of a wider sample of knowledge users, comments by web users in response to the posts made during the campaign across Facebook, Twitter, Instagram, and LinkedIn were collected manually and subjected to a coding reliability thematic analysis, consistent with the method outlined in the Semistructured Interviews With Knowledge Users section. The aim was to understand the type of comment (ie, opinion, personal experiences, and information) in response to the content shared to determine prevalent themes.

Overview of Evaluation Questionnaires

A brief survey was circulated across the campaign’s social media platforms to receive feedback about users’ perception of the content and information shared during the initiative. The participants were recruited via invitations on social media platforms. For example, social media posts were created to encourage participants to provide their input on the initiative by clicking on the link that directed users to the questionnaires. The link to the questionnaire was also available on the website for this project. Individuals who accessed the surveys were...
asked to indicate the group they belonged to (1) caregiver, family member, or friend of a person with dementia, (2) person living with dementia, (3) public, (4) health professional, or (5) policy maker. Demographic information questions about participant age, gender, country, relationship with the person living with dementia, and type of health care provider (ie, for individuals who identified as a health care professional) were collected and outlined using descriptive and frequency statistics.

Pain in Dementia Evaluation Questionnaire
A short questionnaire, used in the pilot campaign evaluation [9], was used to solicit anonymous responses from various knowledge users. The survey included both general and specific questions for the general public, health professionals, and caregivers and family of people with dementia. Most responses were provided along 7-point Likert scales (eg, 1=not very likely to 7=very likely) and focused on viewers’ opinions about the initiative, content, and likelihood of using the information provided. Descriptive statistics were calculated for the Likert scale items in the evaluation questionnaire.

Information Assessment Method for All
Users’ perception of the content and potential benefit of the information were assessed using the IAM4all (Information Assessment Method for All) [30]. The IAM4all consists of 6 general questions assessing information relevance and use, with nested follow-up questions depending on participant responses. Accordingly, the IAM4all is a 28-item self-report questionnaire designed to measure 4 levels of outcomes associated with receiving or seeking web-based information: situational relevance, cognitive impact, information use, and health benefits. Each item is rated based on agreement to a question (ie, yes, no, or possibly). Content validity of the IAM4all has been substantiated through a review of the theories of information science and feedback from knowledge users (eg, laypersons, health professionals, and experts) [30]. Descriptive and frequency statistics were used to calculate the means and SDs for the items in the questionnaire.

Web-Based Discussions About Pain in Dementia
The level of web-based discussions about pain in dementia was determined by examining the number of posts on Twitter and Facebook. Three periods of comparison were established: (1) a precampaign period (ie, a 7-month period following the end of the pilot campaign and before the start of #SeePainMoreClearly phase II, March 1, 2020, to September 30, 2020); (2) a campaign period (October 1, 2020, to September 30, 2021); and (3) a postcampaign period (ie, a corresponding 7-month period following the end of #SeePainMoreClearly phase II, October 1, 2021, to April 30, 2022). Searches for unique (ie, not retweets or reposts) tweets on Twitter containing the hashtags “#pain #dementia” AND/OR key words “pain dementia” were conducted for all 3 periods using the Keyhole software. A catalog of all the tweets during the specified periods was collated on a spreadsheet by Keyhole. A similar method could not be used to obtain Facebook posts on Keyhole (ie, this feature was not available on Keyhole for Facebook posts); hence, a manual search using the aforementioned keywords was conducted for all 3 periods. That is, the specific keywords were typed in the search bar on Facebook, and the period (eg, precampaign period, campaign period, and postcampaign period) was specified in the search filter. Posts for each platform across the 3 periods were cataloged. The following types of posts were excluded: (1) pertaining to pain as a metaphor (ie, using “pain” as a metaphor for psychological distress, such as in the context of grief “I feel your pain”), (2) unrelated to the problem of pain in dementia, (3) shared by our own research group, and (4) not in the English language.

Results

Engagement and Reach of the Initiative

Web Analytics
Over 60,000 users from 82 countries viewed the web-based repository website, which resulted in 83,247 page views over the 12-month period (ie, October 1, 2020, to September 30, 2021). In examining the manner (ie, how people found the website) in which users were directed to the website, a majority of users (ie, 51,693 users) were directed to the website through links (ie, posts including a link to the website) from social media platforms. Other users directly typed the website link (ie, 6927 users) or searched the website link (ie, 2152). A small number of website users (ie, 116 users) were referred by other sites (ie, partnering organizations linking the web-based repository website on their web page). The blog web pages were viewed 59,919 times. Of the blogs posted during the project, the article written (on the request of our team) by an acclaimed author, Louise Penny, in which she relayed her experiences caring for her husband who lived with dementia [31], garnered most of the views (ie, received 33,226 views).

Social Media Analytics
The analytics are summarized in Table 1 based on the 3 sources of information and data periods described in the Methods section. The advertisements that were initiated on Facebook could also be posted on Instagram based on advertisement specifications. To maximize reach, the digital media team specified advertisements on Facebook and Instagram to be displayed on all available placements and to the specified targeted audiences. This includes Facebook and Instagram feeds (ie, advertisements displayed on the user’s feed), in-stream advertisements for videos, search results, and external apps and sites (ie, advertisements displayed to audiences on additional websites and mobile apps, such as newspaper websites). That is, advertisements initiated on Facebook could be displayed on Facebook, Instagram, or to external websites or apps connected to Facebook, resulting in increased impression and reach than organically derived engagement. The advertisements analytics (eg, impression and reach) on Facebook and Instagram differs from Twitter and LinkedIn as a function of the data source (ie, paid impression or reach on Facebook and Instagram are higher because they capture not just impression or reach derived on Facebook and Instagram but also on third-party websites such as newspaper websites; Table 1).

https://aging.jmir.org/2024/1/e53025
Table 1. Summary of the social media analytics for the #SeePainMoreClearly campaign.

<table>
<thead>
<tr>
<th>Period, metric, and source of information&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Facebook, n</th>
<th>Twitter, n</th>
<th>Instagram, n</th>
<th>LinkedIn, n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Period 1&lt;sup&gt;b&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impressions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facebook overall analytics&lt;sup&gt;c&lt;/sup&gt;</td>
<td>4,100,000</td>
<td>—&lt;sup&gt;d&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Twitter only analytics&lt;sup&gt;e&lt;/sup&gt;</td>
<td>—</td>
<td>724,200</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Instagram only analytics&lt;sup&gt;e&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
<td>4883&lt;sup&gt;f&lt;/sup&gt;</td>
<td>—</td>
</tr>
<tr>
<td>LinkedIn advertisement analytics&lt;sup&gt;g&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>8538</td>
</tr>
<tr>
<td>Twitter advertisement analytics&lt;sup&gt;g&lt;/sup&gt;</td>
<td>—</td>
<td>365,935</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Reach</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Instagram advertisement analytics&lt;sup&gt;h&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
<td>138,833</td>
<td>—</td>
</tr>
<tr>
<td>Instagram only analytics&lt;sup&gt;e&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
<td>4420</td>
<td>—</td>
</tr>
<tr>
<td>Facebook only analytics&lt;sup&gt;i&lt;/sup&gt;</td>
<td>1,313,485</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Engagement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facebook only analytics&lt;sup&gt;i&lt;/sup&gt;</td>
<td>282,704</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Cloud Campaign&lt;sup&gt;j&lt;/sup&gt;</td>
<td>—</td>
<td>—</td>
<td>977</td>
<td>—</td>
</tr>
<tr>
<td><strong>Period 2&lt;sup&gt;k&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impressions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facebook only analytics&lt;sup&gt;i&lt;/sup&gt;</td>
<td>519,929</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Reach</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Engagement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Twitter only analytics&lt;sup&gt;e&lt;/sup&gt;</td>
<td>—</td>
<td>630</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td><strong>Period 3&lt;sup&gt;l&lt;/sup&gt;</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Impressions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cloud Campaign&lt;sup&gt;j&lt;/sup&gt;</td>
<td>871,800</td>
<td>—</td>
<td>—</td>
<td>1267</td>
</tr>
<tr>
<td>Reach</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Engagement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cloud Campaign&lt;sup&gt;j&lt;/sup&gt;</td>
<td>—</td>
<td>2125</td>
<td>—</td>
<td>193</td>
</tr>
</tbody>
</table>

<sup>a</sup>Some analytics could not be obtained for specific platforms because they were not available. Moreover, the digital marketing team could only obtain data for specific periods (eg, year-long data could not be obtained) for some platforms owing to limitations in the sources of information. Nonetheless, efforts were made to capture the entire period.

<sup>b</sup>Period 1=October 1, 2020, to September 30, 2021.

<sup>c</sup>“Facebook overall analytics”=generated from paid and unpaid impressions on Facebook and third-party websites such as newspaper websites that are part of the Facebook advertising display network across Canada.

<sup>d</sup>Data for this period or metric were not applicable or not available.

<sup>e</sup>“Twitter/Instagram/LinkedIn only analytics”=analytics based on unpaid posts circulated only on Twitter, Instagram, or LinkedIn.

<sup>f</sup>On Instagram, the period 1 impressions only include impressions for unpaid posts posted on Instagram. It does not include paid posts that were posted on Facebook but reached Instagram audiences.

<sup>g</sup>“Twitter/LinkedIn advertisement analytics”=analytics based on paid posts circulated on Twitter and LinkedIn.

<sup>h</sup>“Instagram advertisement analytics”=the analytics are based on paid posts that were posted on Facebook but reached Instagram audiences based on specified target audiences.

<sup>i</sup>“Facebook only analytics”=based on paid and unpaid posts that were circulated only on Facebook.

<sup>j</sup>CloudCampaign (ie, third-party social media monitoring software): these analytics are based on unpaid posts circulated individually on Twitter, LinkedIn, and Instagram. The Cloud Campaign analytics from Facebook includes analytics of paid and unpaid posts circulated only within Facebook.

<sup>k</sup>Period 2=October 1, 2020, to November 2, 2020.

<sup>l</sup>On Instagram, the period 3 impressions only include impressions for unpaid posts posted on Instagram. It does not include paid posts that were posted on Facebook but reached Instagram audiences.
For a graphical illustration of the analytics, refer to Multimedia Appendix 3. Social media analytics indicated that 1,313,485 people were reached by the content shared on the Facebook page created for this project (Table 1). The #SeePainMoreClearly hashtag on Twitter (eg, obtained by Keyhole) was used by 540 users, resulting in 2835 posts. Moreover, the hashtag reached 1,691,440 users and gained 8,592,929 impressions and 8696 engagements. The 4 animated videos developed during the initiative were viewed over 257,000 times across the web page and social media platforms. The views for each video were as follows: 106,055 (family), 102,407 (health care professionals), 42,731 (policy makers), and 5911 (researchers). The pilot campaign video was also disseminated during phase II of the campaign across social media platforms and received 45,637 views (ie, 43,870 views on Facebook, 1043 on Twitter, 709 views on YouTube, and 15 views on Instagram).

Knowledge Users’ Responses to the Initiative

Semistructured Interviews

Overview

A total of 13 knowledge users who interacted with the initiative’s posts participated in semistructured interviews. All participants provided informed consent. Participants had varied perspectives: 1 participant had lived experience of dementia, 4 participants had family members living with dementia and were health professionals, 4 participants were caregivers, and 4 were clinicians. A thematic analysis was conducted by 2 separate coders to establish an intercoder agreement. A randomly selected 20% (3/13) subset was coded by a second coder to assess for agreement. Overall agreement was excellent (k=0.81). Five themes emerged from the data: (1) increased awareness about the problem of pain in dementia, (2) perceived barriers to pain management, (3) increased knowledge and changes in behavior, (4) value of social media as a method of scientific dissemination, and (5) suggestions for improvement.

Increased Awareness About the Problem of Pain in Dementia

A majority of participants expressed positive views about the initiative, particularly about the manner in which the content was shared throughout the initiative and the resulting increased awareness about the problem of pain in dementia. One participant noted the importance of the information shared during the initiative:

> Overall as a whole, I really liked it. I like the messaging that comes out of it, and I think it’s something really important to continue doing...I don’t think there’s a real understanding about pain and older adults. I like the messaging and I think it’s something that’s really important for us, as well as a society to promote.

Participants also expressed that the information shared during the initiative highlighted the experiences of older adults with dementia and pain. A person indicated that they thought the initiative was “great” because “the more information that can get out to caregivers, the better it will be for the residents [and] the people that have loved ones at home.” Furthermore, the initiative brought awareness to an underdiscussed problem of pain in dementia. For example, one participant stated:

> Oh it’s good thing to put out here, something that people haven’t really thought about even the health care system and [pain in dementia] is a huge problem...so to bring it to the forefront is good, it needs to be brought to the forefront.

Overall, participant responses described a lack of prior understanding related to the experience of pain in dementia and the need for increased awareness about this topic. Some individuals commented on the significance of the specific messages shared throughout the campaign in raising awareness about pain in dementia. For example, a participant noted the following:

> I love what you’re doing. It’s a concern that everybody has, and why I call you a niche [because] it is a subject that people haven’t traditionally thought of.

Similarly, one participant expressed that the messaging on the videos were informational. This participant noted the following:

> I guess one thing that stood out to me so much was the video of the man who seemed very angry. And so, I think back of all the people who’ve been labeled as the “angry person” could be the person in pain. And I think that message really hit me the most on pain.

Several participants expressed that the messaging shared during the initiative changed their perception about the problem of pain in people with dementia. A person commented on the videos:

> You know one thing I did really like is the video with the gentleman that they thought was being kind of aggressive...I thought that video was really good but then on the other side [you] have the gentleman communicating how he is feeling. That was really good because that was really eye-opening.

Perceived Barriers to Pain Management

Knowledge users described various barriers that they have experienced in managing pain in their clinical practice or as a family member caring for a person living with dementia. The barriers encountered by knowledge users centered on inadequate pain assessments conducted in their practice or of their family member and relative living with dementia. For example, a health care professional reported the inadequate frequency of pain assessment in their facility:

> You know, often in Ontario anyway, where I live, assessments in long-term care are done quarterly on clients and that’s, you know, that’ll be an all-encompassing assessment but certainly pain is part of that evaluation and, I mean, four times a year is not even close to being adequate to properly address and intervene if someone is living with pain.

In light of inadequate pain assessment practices, caregivers are often left to advocate for their family member or relative living...
with dementia. A caregiver in the study expressed the importance of families in advocating for assessment for their loved one:

_I realized I don’t know if [my mother] might have tooth decay in her teeth [or if it] would be hurting her, and I don’t know how to figure that out. And I don’t think anyone else is checking. Yeah, because I think the lesson of my story here is that I am more likely to assess pain in my mom than her [care providers]_

Participants described lack of continuing education and awareness as significant barriers to improving practices in LTC settings. For instance, a participant indicated the following:

_I think there’s certainly a lack of understanding and education by many health care professionals who are working with older adults who live with dementia, so I think a barrier is getting that education, well, I mean, it’s an opportunity for people to have more education but a barrier is kind of reaching the people._

Participants also indicated that the lack of education of the public at large was a significant barrier to improving current practices:

_I think the barrier is the lack of education, the lack of knowledge that people with dementia can have pain and express it in a different way. So, I think that the lack of education of health care workers, but the public in general._

Insufficient education about the topic means that health professionals or caregivers of persons with dementia may not be aware of the signs and symptoms that can indicate pain. A caregiver in the study reflected on this notion and shared their lack of knowledge about the signs of pain in her mother living with dementia:

_I mean I can think back to my mother who had two strokes. And after the second one, she used to scream all the time. She used to scream. And she never spoke to anyone, but she screamed. And when I questioned the doctor, if could she be in pain, he was so sure that she was not. It doesn’t manifest that way. And so, I think back like, “was she in pain all the time?” When she screamed louder when we came, was it a plea for help. I think lot about that._

**Increased Knowledge and Changes in Behavior**

Participants described the positive impact of the information on their knowledge and behavior. Some participants noted that the information shared during the initiative influenced their awareness and advocacy in their personal life or in clinical practice. Health professionals indicated that they found gaps in their own practice as a result of viewing the content shared in this initiative. A participant said the following:

_It made me more aware of the gap in pain assessment with people like my mom. And then it also made me more aware that the onus then is on me as a caregiver to be doing it, and I feel woefully inadequate, that fact that I encountered the research has made me realize that it’s me doing the diagnostic [work] and I have no medical skill at all._

Greater insight into their own behavior was also reported by the participants. A clinician noted the following:

_It also made me more aware for myself in my own practice...so that’s changed my practice in that way. To be more cognizant that behaviors the patient is displaying could be manifested due to pain. So, I need to ask the caregivers and ask the family, “Has he or she changed in behaviors in any way recently?” “Have you noticed, you know, a consistent type of behavior? Is it a specific time of day, is it related to something?” So that I could look more into pain and assessing it in a roundabout way, rather than asking them straight. “Do you think that your husband or wife is in pain?”_

Other participants indicated that the initiative largely impacted their perspective on this issue (eg, “I would say it’s definitely changed the viewpoint”). Finally, participants indicated that they shared the knowledge they obtained with others:

_I share things that you guys put out there, if you look back at my timeline, you’ll see periodically I make reference to you guys._

Similarly, one participant noted the following:

_I forward all this information off to my colleagues that work in this [area] that are even more focused in the nursing aspect that would have a more direct impact on client care needs._

**Value of Social Media as a Tool for Scientific Dissemination**

Participants underscored the importance of leveraging social media to share research information in the age of the internet. Participants expressed that sharing evidence-based information fosters trust and credibility among users seeking information on social media platforms. For example, one participant noted the following:

_I think it has a lot of positive benefits. Where to start? After this past year right, social media and technology were probably utilized more than they ever have been. So, I actually think it’s a really great way to disseminate that information. And then further to that too, because I’ve been following all your accounts for a while, just there again, yeah, the quality of the content is also really good. So, with that being said, it creates that credibility and that trust. And then that also really feeds into where social media sometimes isn’t the most trusted source. So, I do think that it’s been a really great way to get that information out there._

Participants also acknowledged the salience of social media platforms that needs to be further leveraged for scientific dissemination (eg, “social media certainly in light of COVID, like it is the way that information gets disseminated right now so I think it’s very appropriate and it is a strength that you’re on various platforms sharing resources”). Other participants
expressed the importance of health initiatives in combating misinformation over the web:

- I think a lot of times there’s a lot of misinformation for people and families and health care practitioners and [they] want to have some valid research-based information for best practices and when you do things like this and have it more accessible then it’s easier for people to gain the right information and make the best decisions for their family members or their clients that they can.

Other participants commented on the reach that social media platforms hold in raising awareness of issues in a short amount of time and connecting knowledge users across the world. One participant indicated the following:

- It’s brilliant, social media has always been very, very good at distributing information. Creating awareness. Highlighting anything in social media catches the eye of the person. Because people are looking for answers.

In particular, one participant commented on the utility of Facebook targeting specific demographic groups:

- I think, using social media is great, you know, the thing about Facebook, is that it is now an older person’s social media and people often think that that young people are caregivers of people with dementia. But it’s not always so—it’s old people who are caring for people with dementia. Facebook is good because most of us use Facebook.

Suggestions for Improvement
Finally, participants made suggestions about scaling and improving the initiative. One participant said the following:

- I think it would be effective for you to reach out to the Alzheimer’s Society...I would love to see you partner with them in an active way on their media.

One participant suggested focusing on reaching individuals who may not be directly reached by the initiative. For example, one participant noted individuals who may not be using or cannot access social media platforms:

- I think smaller communities don’t have the access or don’t know about it. I mean now we do have the Internet and that sort of thing which is fabulous but lots of time we don’t even know it’s out there.

As another example, one participant stated the following:

- It’s the reachability...say like you know, even for me, it was from [someone else] that I found out about this, I have never seen it while using any of the social media account before. And like that largely depends on my browsing habits too right? But it’s the reachability that you guys have to concentrate on.

Other participants suggested incorporating information to traditionally delivered pain education:

- And so, if there was a way, we could get this added into a pain curriculum, I think that would be excellent, because I never had it in my course, and I’ve learned from this.

In clinical practice, one participant expressed the following:

- Well, I think, for so many people with dementia, their care is being provided by people with minimal education or I should say varying levels. If [people could get] a little certificate put into their little portfolios that would show, they have completed a course in pain assessment for people with dementia and that would somehow be an incentive for them as workers.

Social Media Comments
Overview
A total of 895 comments were included in the analysis. Most of the comments were retrieved from Facebook (eg, 822/895, 91.8%), followed by Twitter (eg, 68/895, 7.6%), and a small subset was comments by users on Instagram (eg, 5/895, 0.6%). A randomly selected 20% (179/895) subset was coded by a second coder to assess for agreement. The overall intercoder agreement was excellent (ie, k=0.80). Six themes emerged from the analysis: (1) positive comments in response to the initiative, (2) sharing their personal experiences in response to the content, (3) criticisms about and suggestions to improve pain management practices, (4) responses related to the COVID-19 pandemic, (5) negative comments in response to the content, and (6) advertisements.

Positive Comments in Response to the Initiative
Many users provided positive comments in response to the content. Facebook users expressed empathy in response to the content shared in the form of comments (eg, “I am praying for you” and “Amen”) and with the use of “emojis” (eg, praying emoji and red hearts). Other respondents commented on the importance of the initiative. For instance, a user on Facebook provided a comment noting the following:

- The work you are doing is so important to all.

As another example, other users commended the initiative (eg, “such an important campaign #seepainmoreclearly”). Blog posts written by care partners in which they relayed their experiences as caregivers stimulated various positive discussions about the initiative. For example, a user commented “thanks for bravely sharing your experience with having a spouse with dementia, what a difficult journey, I enjoy [your books] so much!” in response to Louise Penny’s blog post shared on social media. As another example, a user expressed the following:

- Thank you for this information which is truly important. Your story is important to open our minds to an area of Alzheimer’s that many did not think of.

Sharing Their Personal Experiences in Response to the Content
Other users responded to the initiative by sharing their personal experiences with the posted content. For example, one user stated the following:

- My family went through this when my dad was only in his early 50s...no one had heard of
Alzheimer’s...this disease is so sad and robs the family of so much. I’m glad there’s more support for families now, and wish we’d had more support when we needed it.

Another user expressed their experience of caring for their mother:

I wonder this all the time. My mom is nearly 101 and has dementia. She rarely expressed having pain. But how can she not when she is quite hunched over and has arthritis.

Other individuals provided additional commentary on the shared information. A user emphasized the importance of pain assessment:

People with dementia, people unable to clearly explain [their] pain, we need to look for clues daily that would help us determine if something out of the norm is going on.

Criticisms About and Suggestions to Improve Pain Management Practices

Critical comments about current practices or advocating for improved practices were also present. For instance, a user expressed the following in response to a post:

There is a problem with doctors overlooking any kind of pain.

Another user pointed out the issue of resource limitations in LTC facilities (eg, “staffing is a huge issue which needs a timely resolution”). Moreover, a user indicated that “seniors deserve much more than the less of minimum care they get.” Other users provided suggestions to improve practices:

There are so many kinds of dementia, each with their own stages. We need more access to good education and support for home care to help families cope.

Another user stated the following:

Pain assessment only on admission and then every 3 months? Pain must be assessed whenever there is the slightest indication of pain. A formal assessment every 3 months should reflect how effective the measures were.

Responses Related to the COVID-19 Pandemic

As the campaign occurred during the COVID-19 pandemic and disseminated vaccination information related to residents in LTC, a subset of the comments was related to the pandemic. Some users noted the negative ramifications of isolation in LTC (eg, “I am sick of the lockdowns in care homes. Not being able to take them out for a day for visits...is absolutely insane. They are suffering terribly over this, and it has to stop.”) Other comments denounced the significance of vaccines (eg, “the fake pandemic was created for the COVID vaccines”).

Negative Comments in Response to the Content

A subset of the comments expressed negative responses to the information shared. A few of these comments perpetuated stereotypes about older adults (eg, “people with dementia still feel pain?”) or dementia (eg, “reading this is enough to get dementia”). A user expressed disagreement with one of the information shared:

We assess residents for pain every time we see them.
We are not stupid; we can tell when someone is in pain.

Do any of you work in LTC facility? Staff in dementia units are trained to look for any behavioral issues that may arise from pain.

Advertisements

Finally, a small subset of comments was from users advertising a product or information (eg, “Dementia and Alzheimer’s affects so many. Please don’t forget to check out my podcast, if you haven’t already” and “Can we help your loved one/care home/hospital/nursing home in UK with our free mp3s preloaded with the music of their own choice?? Contact me asap please click on the link.”).

Responses to the Evaluation Questionnaires

The demographic characteristics of the knowledge users who responded to the questionnaires are presented in Multimedia Appendix 4. The majority of the survey respondents were caregivers of people living with dementia. Moreover, most respondents indicated that they had found the #SeePainMoreClearly campaign on Facebook and were living in Canada. Not all participants who responded to the survey completed all the questionnaires; therefore, the number of respondents is indicated in Tables 2 and 3. Table 2 outlines the respondents’ impression of the campaign. Across respondent groups, respondents endorsed a favorable impression of the campaign, a great likelihood of sharing information with others, and the use of social media for KM. Of note, the question regarding the new information provided by the content was rated lowest in each respondent group. Table 3 outlines the descriptive statistics based on the responses to the 6 general questions assessed by the IAM4all questionnaire. As noted in Table 3, responses indicated that most individuals (eg, 138/178, 77.5%) indicated the likelihood of using the information for themselves or someone else. Moreover, 77.1% (131/170) of the individuals expressed benefiting from the information shared during the initiative.
Table 2. Respondents’ impression of the campaign.

<table>
<thead>
<tr>
<th></th>
<th>Caregivers</th>
<th>Public</th>
<th>Health professional</th>
<th>Person living with dementia</th>
<th>Researcher</th>
<th>Policy maker</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall impression of the campaign</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value, mean (SD)</td>
<td>4.67 (1.63)</td>
<td>4.33 (1.49)</td>
<td>5.16 (1.72)</td>
<td>4.18 (1.72)</td>
<td>__ b</td>
<td>—</td>
</tr>
<tr>
<td>Value, n</td>
<td>78</td>
<td>24</td>
<td>25</td>
<td>11</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Level of new information provided</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value, mean (SD)</td>
<td>3.91 (1.83)</td>
<td>4.21 (1.74)</td>
<td>4.04 (2.09)</td>
<td>3.50 (2.07)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Value, n</td>
<td>79</td>
<td>24</td>
<td>25</td>
<td>10</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Impact on information-seeking behavior</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value, mean (SD)</td>
<td>4.90 (1.89)</td>
<td>4.33 (1.76)</td>
<td>4.84 (1.82)</td>
<td>5.10 (2.33)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Value, n</td>
<td>79</td>
<td>24</td>
<td>25</td>
<td>10</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Using social media for scientific dissemination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Value, mean (SD)</td>
<td>5.48 (1.74)</td>
<td>5.00 (1.67)</td>
<td>5.44 (1.53)</td>
<td>5.00 (1.83)</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Value, n</td>
<td>87</td>
<td>25</td>
<td>25</td>
<td>10</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

a Ratings are based on a 7-point scale, ranging from 1 (not at all) to 7 (very much): “Overall, how much did you like the image/clip/video?”
bCell sizes <5 were not reported to protect the confidentiality of the participants.
cn=not all participants who responded to the survey completed all the questionnaires in the study; this number represents the number of participants who answered this question.
d“Did the video/image/clip provide you with new information?”
e“Are you likely to seek additional information about pain in dementia (and its assessment) as a result of watching this video/viewing this content?”
f“It takes an average of 17 years until professionals start to use important research results into their practice Do you think social media (e.g., Twitter) is a good way to spread important health information to caregivers of people with dementia/to people with dementia?”
Table 3. Survey responses to the IAM4all<sup>a</sup> questionnaire (N=190).

<table>
<thead>
<tr>
<th>Question</th>
<th>Respondents, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot;Why did you look for this information?&quot;&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>To answer a question about the health of someone else (total number of respondents=190)</td>
<td>116 (61.1)</td>
</tr>
<tr>
<td>To satisfy my curiosity about a health matter (total number of respondents=172)</td>
<td>108 (62.8)</td>
</tr>
<tr>
<td>To find choices different from those given by a health professional (total number of respondents=172)</td>
<td>61 (35.5)</td>
</tr>
<tr>
<td>To follow-up on the information given by a health professional (total number of respondents=168)</td>
<td>52 (31)</td>
</tr>
<tr>
<td>To prepare myself before talking to a health professional (total number of respondents=168)</td>
<td>50 (29.8)</td>
</tr>
<tr>
<td>To help me decide if I should see a health professional (total number of respondents=169)</td>
<td>40 (23.7)</td>
</tr>
<tr>
<td>To answer a question about my health (total number of respondents=177)</td>
<td>30 (17)</td>
</tr>
<tr>
<td>&quot;Did you find the information you were looking for?&quot; (total number of respondents=181)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>73 (40.3)</td>
</tr>
<tr>
<td>Yes, but I did not understand it</td>
<td>15 (8.3)</td>
</tr>
<tr>
<td>No, I did not find it</td>
<td>72 (39.8)</td>
</tr>
<tr>
<td>No, but I found something else</td>
<td>21 (11.6)</td>
</tr>
<tr>
<td>&quot;What did you think about this information?&quot;&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
</tr>
<tr>
<td>Now I want to learn more about this health matter (total number of respondents=161)</td>
<td>112 (69.6)</td>
</tr>
<tr>
<td>Now I know something new (total number of respondents=163)</td>
<td>79 (48.5)</td>
</tr>
<tr>
<td>I am reminded of something I already knew (total number of respondents=157)</td>
<td>76 (48.4)</td>
</tr>
<tr>
<td>This information says I did, or I am doing the right thing (total number of respondents=158)</td>
<td>66 (41.8)</td>
</tr>
<tr>
<td>Now I am reassured (total number of respondents=158)</td>
<td>66 (41.8)</td>
</tr>
<tr>
<td>I am not satisfied with this information (total number of respondents=153)</td>
<td>33 (21.6)</td>
</tr>
<tr>
<td>I think there is a problem with this information (total number of respondents=151)</td>
<td>14 (9.3)</td>
</tr>
<tr>
<td>I think this information could be harmful (total number of respondents=148)</td>
<td>2 (1.4)</td>
</tr>
<tr>
<td>&quot;Did you or will you use the information for yourself&quot; (total number of respondents=178)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>55 (30.9)</td>
</tr>
<tr>
<td>No, not for myself, but I used it for someone else</td>
<td>83 (46.6)</td>
</tr>
<tr>
<td>No, I did not use this information for myself or for someone else</td>
<td>40 (22.5)</td>
</tr>
<tr>
<td>&quot;Did you (do you expect to) benefit from this information?&quot;&lt;sup&gt;b&lt;/sup&gt; (total number of respondents=170)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>131 (77.1)</td>
</tr>
<tr>
<td>No</td>
<td>39 (22.9)</td>
</tr>
<tr>
<td>Did something negative come out from using this information? (total number of respondents=148)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>4 (2.7)</td>
</tr>
<tr>
<td>No</td>
<td>144 (97.3)</td>
</tr>
</tbody>
</table>

<sup>a</sup>IAM4all: Information Assessment Method for all.

<sup>b</sup>Refer to participants who answered “Yes” with possible choices of yes, no, or possibly.

Comparison of Web-Based Discussions About Pain in Dementia

The extent of discussions about pain in dementia was assessed by examining the total number of posts on Twitter and Facebook. After excluding posts pertaining to pain as a metaphor, unrelated to the problem of pain in dementia, shared by our own research group, and not in the English language, the number of tweets that were retained doubled from the precampaign period compared with during the campaign period (see Multimedia Appendix 5). A similar increase was also observed on Facebook. The number of discussions about pain in dementia decreased in the postcampaign period compared with during the campaign period (see Multimedia Appendix 5). The number of excluded posts in each period is summarized in Multimedia Appendix 6.
Discussion

Principal Findings
Social media allows the immediate dissemination of information to a large number of knowledge users. Although our relatively short pilot KM campaign (phase I) was successful, its longer-term continuation by researchers alone was not feasible owing to competing demands. As such, we partnered with a digital media company to launch phase II of the #SeePainMoreClearly campaign with expanded social media platform coverage. The goal of the initiative was to increase awareness of the challenges in the assessment and management of pain among people with dementia. We codified key messages with partners (eg, health professionals, caregivers, and researchers) that served as the basis of our messaging.

Our findings demonstrated the effectiveness of social media KM methods in reaching very broad audiences over a 12-month period. Evidence of positive impact on knowledge users’ knowledge was demonstrated. The participants described positive perceptions in response to the information shared throughout the initiative. Survey respondents (eg, caregivers, health professionals, researchers, the public, and people with dementia) endorsed a favorable impression of the campaign. Overall, our initiative highlighted the advantages of using a science-media partnership (eg, collaboration with digital media experts, resources to develop tailored content and resources, and consistent dissemination of information). However, broad social campaigns require extensive resources and time commitments. We outline the recommendations to address these areas.

The current (phase II) initiative involving a science-media partnership was different in several regards from our pilot campaign that focused only on Twitter and a YouTube video. In comparison with the pilot campaign [9], visits to our web-based repository substantially increased in phase II, even when considering the difference in the lengths of the 2 campaigns. Phase II used extended evaluation metrics through the inclusion of interviews, in addition to social media analytics, questionnaires, and the analysis of social media responses over a 12-month period. Hashtag analytics on Twitter across the 2 phases were comparable, despite differences in the length of evaluation (eg, 5 months vs 12 months). We observed lower outside organization engagement during phase II of our initiative, which could have contributed to lower phase II engagement when accounting for the length of evaluation. Moreover, our social media presence was greater in phase II compared with our pilot campaign.

Overall, our findings demonstrated the effectiveness of social media KM methods in reaching very broad audiences quickly. In this study, we demonstrated success in directing knowledge users to a resource website with practical information for health professionals, caregivers, and people living with dementia. For instance, over 60,000 users from 82 countries viewed the web-based repository website over the 12-month period. In particular, the blog posts shared during the project attracted many readers. The blog web pages were viewed 59,919 times, providing evidence for their reach. This is comparable with other KM efforts showing the use of blogs as an engaging way of connecting to targeted users [32].

The reach and engagement of users on social media platforms also provided evidence for the success of the initiative in raising awareness about the problem of pain in dementia. Information shared on Facebook was successful at reaching a wide range of audience, with 1,313,485 people reached by the content shared on the Facebook page. However, it should be noted that some analytics (eg, reach and engagement) could not be obtained from social media sites (eg, LinkedIn) because these analytics were not provided by the platform. This limits our ability to deduce the full extent of engagement on these platforms. Nonetheless, the number of impressions on Facebook was the highest (eg, 4,100,000) compared with Twitter, Instagram, and LinkedIn. This is consistent with previous research showing extensive reach and engagement on Facebook compared with other platforms [6,33]. In contrast, Neil-Sztramko et al [34] conducted an awareness campaign targeted toward working caregivers and found that although Facebook posts generated the most reach, the quality of the engagement was low. The topic area, targeted audience, and length of evaluation could explain the differences between our findings and those of previous research. We observed the lowest reach and engagement on Instagram and LinkedIn. This is not surprising given that our primary audience comprised caregivers of people with dementia who tend to be older [35,36]. Older adults have shown a preference for Facebook in comparison with other social platforms [37]. Taken together, our findings suggest that Facebook may be the most suitable platform for disseminating information related to pain in dementia.

Our findings extend our pilot evaluation by including in-depth interviews with knowledge users who interacted with the campaign in addition to the analysis of social media analytics and evaluation questionnaires. Many interview participants indicated positive perceptions in response to the information shared throughout the initiative. Moreover, the participants expressed that the significance of the initiative was in bringing awareness to an underdiscussed problem. The interview participants acknowledged the utility of social media as a tool for scientific dissemination. Many participants noted the importance of leveraging social media to share research information. Other participants expressed the importance of health initiatives in combating misinformation over the web. Quantitative responses to the surveys also supported this view. Survey respondents endorsed the use of social media as a way of spreading important health information to caregivers of people with dementia and to people with dementia. We found evidence of the information and resources impacting users’ knowledge and behavior. For example, interview participants noted increased awareness and advocacy in their personal life and awareness of assessing for pain in their clinical practice. More importantly, participants frequently expressed intent to share the knowledge they obtained with others. This is consistent with the survey responses, indicating that a majority of respondents intended to use the information for themselves or others. Moreover, the respondents expressed benefitting from the information shared during the initiative.
The initiative stimulated web-based discussions about pain in dementia. Consistent with themes that emerged from social media responses on Twitter during the pilot campaign [9], many web users expressed positive comments in response to the initiative. In particular, expressions of support were prevalent in both the pilot and phase II of our campaign. Many web users who responded to the content disseminated during the campaign shared their personal experiences or added commentary to the posted content. In addition to sentiments of advocacy for better care, which was observed in our pilot campaign, social media comments in response to phase II of our campaign also highlighted criticism and suggestions to improve practices related to pain management in dementia. Users highlighted the need for access to continuing education and support for staff and families to increase the frequency of pain assessment in LTC. This demonstrated meaningful engagement by users in response to posts. Increased discussions also highlighted negative responses about the information shared (eg, perpetuating stereotypes about older adults and dementia), which was not observed in the pilot campaign. Differences in themes that emerged between the pilot and phase II could be due to the broader reach of phase II (eg, inclusion of Facebook), as the pilot campaign only examined responses using the #SeePainMoreClearly hashtag on Twitter. We also found increased discussion about pain in dementia during the initiative on Twitter and Facebook in comparison with before and after the initiative. However, this discussion was not sustained months following the end of the campaign. This highlights the importance of continuous and meaningful engagement to maintain gains made by an initiative. Given our methodology, we collected a substantially lower number of posts on Facebook in comparison with Twitter. This discrepancy is likely owing to the manual search that was conducted to obtain Facebook posts in comparison with the data that were easily pulled by Keyhole for Twitter posts. Depending on a user’s privacy settings, posts shared on personal or private accounts are excluded when conducting a general search on Facebook.

Limitations and Directions for Future Research

A particular strength of this study is that the content shared during the initiative was developed collaboratively with family caregiver partners, researchers, media experts, and health care professionals. Various measures were also used to assess not only the web-based reach of the initiative but also the perceptions of knowledge users. Nonetheless, we acknowledge that the use of social media for KM research represents a new area that does not lend itself to strict scientific control in the evaluation of its effectiveness. We adapted a more nuanced approach in developing partnerships and using multimodal approaches to show the impact in different ways. Although we aimed to assess the impact on users’ knowledge and behavior, we did not directly evaluate behavior change. Future research should examine the associated behavior outcomes (eg, increased pain assessment) in addition to the spread of information (eg, analytics). It is also difficult to measure the impact of the initiative at an organizational or policy level. This could be an important avenue for future research.

In our interview and surveys expressed their intention to use the information; however, the adoption and use of information was not obtained. Perhaps, a follow-up assessment of how or whether knowledge users applied learned information could be investigated. Moreover, the effectiveness of social and web modalities in creating meaningful changes in knowledge and behavior is subject to further examination. Notably, the landscape of social media is constantly changing, and researchers will need to adapt to these changes if they want to leverage this modality. For example, social media can be used to spread health misinformation [38]. Using social media for KM health efforts can also be used to combat false and unscientific web-based information. As such, ensuring the quality of scientific evidence disseminated on the web should be considered in the development of social media–based dissemination efforts. Although these findings may not be generalizable to other KM initiatives, our investigation adds to the growing body of knowledge leveraging social media as a KM tool.

Recommendations

Our study highlighted the successes and challenges of social media KM initiatives. When comparing the pilot with the scaled initiative (phase II), there appears to be no considerable difference with regard to views, impressions, or reach when accounting for the length of time of each initiative (5 months vs 12 months). Although this scaled initiative demonstrated success in garnering a large reach across social media platforms, broad initiatives such as phase II of our campaign require extensive resources and time commitment, which may not be feasible for researchers in the long term. The following recommendations are proposed to assist researchers and partners who may be interested in developing and maintaining a web-based KM initiative:

1. Researchers could collaborate with established communication and marketing departments at their institution or other digital media partners to aid in digital and social media outreach of research information and internal and external advertising. Granting agencies may be able to allocate funds for such purposes.
2. Facebook and Twitter appear to be the most relevant platforms for KM for the topic of pain and dementia. However, researchers from other disciplines may find other platforms more helpful in reaching their target audience. For example, if a researcher’s area of research is adolescent mental health, Instagram or TikTok and the use of stories may increase engagement among knowledge users. Although we found success in using multiple social media platforms, it may be more sustainable to pool resources on 1 or 2 of the most used platforms based on the target audience.
3. Investments in targeted and paid advertisements can be an effective mechanism to increase the exposure and engagement of campaign posts at a low cost per engagement [39,40].
4. Our initiative highlighted the importance of codeveloping the initiative with partners of our intended target groups. Many of our partners developed blog posts about their lived experience, which generated the most engagement. As such, personalized blog posts and opportunities for knowledge users to share their experiences (ie, turning comments on)
can be an easy and cost-effective way to stimulate discussions.

5. The success of our pilot initiative was largely influenced by partnering with professional organizations that helped disseminate information and resources to their audiences. Engagement from external organizations was lower during phase II of our initiative, potentially contributing to the overall decrease in engagement. We observed that external organizations were more inclined to retweet our content when it originated from our individual researcher accounts rather than from a generic SeePainMoreClearly social media account. Researchers should collaborate with organizations within their specific area of interest who can support in meaningfully engaging with their targeted audiences. Connecting with trusted messengers (eg, influencers, people with lived experience, and large followings) could help to penetrate targeted groups.

Conclusions
We partnered with a digital media partner to launch phase II of the #SeePainMoreClearly campaign with expanded social media platform coverage. The goal of the initiative was to increase awareness of and provide resources related to the challenges in the assessment and management of pain among people with dementia. Although this scaled initiative demonstrated success in garnering large reach across social media platforms, broad initiatives such as phase II of our campaign (reported in this manuscript) require extensive resources and time commitments, which may not be feasible for researchers in the long term. Researchers should leverage collaborations with their institutions to aid in the digital media outreach of research information. Furthermore, granting agencies should consider allocating more funds for such KM purposes. Our initiative highlighted the importance of codeveloping KM efforts with the partners of our intended target groups and working with professional organizations to disseminate information to our target audience. Collaborations with people with lived experiences and professional organizations will be key to the success of any future KM effort. Our study adds to the growing body of knowledge that leverages social media as a KM tool.

Acknowledgments
This study was supported by funding from the Saskatchewan Centre for Patient-Oriented Research (#345325) and the Saskatchewan Health Research Foundation (#23629).

Authors' Contributions
TH wrote the grant application that supported this research; played a key role in project conceptualization, knowledge mobilization (KM) messaging, analytic direction, and manuscript writing; and oversaw all aspects of this study. LIRC prepared the first draft of this manuscript and played a key role in project conceptualization, data collection, data analysis, and KM messaging. MB, a partner with lived experience as a caregiver, played a key role in KM messaging; provided input in the manuscript; and contributed to team discussions. VT contributed to data analysis. CTC was a consultant on research design and development, provided detailed feedback on the manuscript, and participated in team discussions about project progress and development. SW, a digital media partner, oversaw the collection of web analytics and digital dissemination. All authors participated in team discussions that directed this work, provided input on multiple aspects of the project (eg, social media dissemination strategy) and messaging content, and reviewed and approved the manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Key points (cross-cutting messages).
[DOCX File, 21 KB - aging_v7i1e53025_app1.docx ]

Multimedia Appendix 2
Interview moderator guide.
[DOCX File, 15 KB - aging_v7i1e53025_app2.docx ]

Multimedia Appendix 3
Summary of the social media analytics for the #SeePainMoreClearly campaign.
[DOCX File, 28 KB - aging_v7i1e53025_app3.docx ]

Multimedia Appendix 4
Demographic characteristics of the survey respondents.
[DOCX File, 15 KB - aging_v7i1e53025_app4.docx ]
Multimedia Appendix 5
Number of posts on Twitter and Facebook about pain in dementia.
[DOCX File, 19 KB - aging_v71e53025_app5.docx]

Multimedia Appendix 6
Number of excluded posts for each criterion.
[DOCX File, 21 KB - aging_v71e53025_app6.docx]

References
10. Chambers CT. From evidence to influence: dissemination and implementation of scientific knowledge for improved pain research and management. Pain 2018 Sep;159 Suppl 1(1):S56-S64. [Medline: 29847349]
34. Neil-Sztramko SE, Dobbins M, Williams A. Evaluation of a knowledge mobilization campaign to promote support for working caregivers in Canada: quantitative evaluation. JMIR Form Res 2023 Jun 22;7:e44226 [FREE Full text] [doi: 10.2196/resprot.4808] [Medline: 37347525]

Abbreviations

IAM4all: Information Assessment Method for all
KM: knowledge mobilization
LTC: long-term care
Social Media Discourse Related to Caregiving for Older Adults Living With Alzheimer Disease and Related Dementias: Computational and Qualitative Study

Andrew C Pickett¹, MSED, PhD; Danny Valdez², PhD; Kelsey L Sinclair², MA; Wesley J Kochell², BSPH; Boone Fowler²; Nicole E Werner¹, PhD

¹Department of Health & Wellness Design, School of Public Health, Indiana University, Bloomington, IN, United States
²Department of Applied Health Science, School of Public Health, Indiana University, Bloomington, IN, United States

Corresponding Author:
Andrew C Pickett, MSED, PhD
Department of Health & Wellness Design
School of Public Health
Indiana University
1025 E 7th St
Bloomington, IN, 47405
United States
Phone: 1 812 855 1561
Email: picketac@iu.edu

Abstract

Background: In the United States, caregivers of people living with Alzheimer disease and Alzheimer disease–related dementias (AD/ADRD) provide >16 billion hours of unpaid care annually. These caregivers experience high levels of stress and burden related to the challenges associated with providing care. Social media is an emerging space for individuals to seek various forms of support.

Objective: We aimed to explore the primary topics of conversation on the social media site Reddit related to AD/ADRD. We then aimed to explore these topics in depth, specifically examining elements of social support and behavioral symptomology discussed by users.

Methods: We first generated an unsupervised topic model from 6563 posts made to 2 dementia-specific subreddit forums (r/Alzheimers and r/dementia). Then, we conducted a manual qualitative content analysis of a random subset of these data to further explore salient themes in the corpus.

Results: The topic model with the highest overall coherence score (0.38) included 10 topics, including caregiver burden, anxiety, support-seeking, and AD/ADRD behavioral symptomology. Qualitative analyses provided added context, wherein users sought emotional and informational support for many aspects of the care experience, including assistance in making key care-related decisions. Users expressed challenging and complex emotions on Reddit, which may be taboo to express in person.

Conclusions: Reddit users seek many different forms of support, including emotional and specific informational support, from others on the internet. Users expressed a variety of concerns, challenges, and behavioral symptoms to manage as part of the care experience. The unique (ie, anonymous and moderated) nature of the forum allowed for a safe space to express emotions free from documented caregiver stigma. Additional support structures are needed to assist caregivers of people living with AD/ADRD.

(JMIR Aging 2024;7:e59294) doi:10.2196/59294

KEYWORDS
caregiving; dementia; social support; social media; Reddit
Introduction

Background

There are an estimated 11 million people in the United States currently living with Alzheimer disease and Alzheimer disease–related dementias (AD/ADRD)—a number that is expected to more than double by 2050 [1]. Most people living with AD/ADRD receive informal (ie, unpaid) care support from family members or other personal connections due to the high costs of professional care, the limited capacity of assisted living and memory care facilities, and a strong desire to age in place [2-5]. Recent estimates suggest informal AD/ADRD caregivers in the United States provide >16 billion hours of care labor annually. This has led to calls from national health organizations to prioritize the expansion of support systems for informal caregivers of people living with AD/ADRD. Caregiving for people living with AD/ADRD is a complex and challenging role that is associated with a broad set of suboptimal economic, physical, and psychological outcomes. For example, caregiver burden is associated with negative mental health outcomes, including high levels of stress and depressive symptomology. However, social support may offer a means for reducing caregiver burden and improving health outcomes for both caregiver and care recipient. This study aimed to understand the ways AD/ADRD caregivers seek and experience social support on the web, specifically through the social media platform Reddit.

Caregiving for Individuals Living With AD/ADRD

Recent studies have estimated that approximately 15 million Americans currently provide care for individuals living with AD/ADRD [6]. Most are classified as informal caregivers, meaning they are not paid for labor and time associated with care responsibilities and are typically family members or friends of the person with AD/ADRD. The average life expectancy for individuals diagnosed with dementia can vary widely, from 3 to 10 years, depending on the specific diagnosis, age at the time of diagnosis, and other health factors [7]. Not surprisingly, the caregiving experience can also vary widely. A recent study found that near the time of diagnosis, people with AD/ADRD received an average of 151 hours (approximately 6.5 days) of caregiving monthly, typically provided by 1 caregiver. However, over time, their needs progress, requiring nearly twice the hours of care and support of additional caregivers [8].

Caregiver burden describes the multifaceted strain faced by individuals in providing care, which may include financial, emotional, and physical stressors [9,10]. The challenges of AD/ADRD care are associated with significant health impacts for caregivers [11,12]. Providing care is often associated with higher levels of reported stress [13]. Furthermore, caregivers report high levels of psychiatric symptoms, most commonly depressive symptoms—particularly as the care recipient’s AD/ADRD symptomology progresses [14,15]. Caregivers also report high levels of financial strain related to high costs of care and lost earning potential due to time commitments of care responsibilities [9,16].

Regardless of the dementia subtype, a broad range of neuropsychiatric symptoms, including both behavioral and cognitive changes, can present [17]. Behavioral symptoms of AD/ADRD, for example, can include confusion, aggression, and increased hospital or emergency department visits. Dementia is also associated with neuropsychiatric symptoms, including agitation, depression, hallucinations, anxiety, and apathy [18-20]. Furthermore, dementia is often associated with reduced function of other organs, resulting in symptoms such as voice or speaking challenges, skin injuries, urinary incontinence, constipation, urinary tract infections, dental and vision problems, and hearing loss, among others [18-20]. Some symptoms, such as anxiety and depression, are fairly common across patients with AD/ADRD and with disease progression [20]. Others, such as psychosis, aggression, and agitation, are often exacerbated as cognitive decline increases and are thus associated with steep increases in caregiver burden due to the impact these symptoms have on the completion of basic activities of daily living [18].

As AD/ADRD progresses, more severe symptoms reduce individuals’ capacity to independently complete activities of daily living, thereby necessitating increased assistance and caregiver supervision [21]. Therefore, the time spent on caregiving activities and the number of individuals providing care generally increase across the disease progression [2,3]. Naturally, increased time spent on care, combined with increasingly severe symptoms to manage, is associated with increased feelings of burden. As such, caregivers often spend less time and energy on their own self-care, particularly as the care recipient’s neuropsychiatric symptoms worsen [22]. The emotional, financial, and physical strain of caregiving may contribute to reduced overall health of caregivers [5]. AD/ADRD caregivers often report high levels of stress and depression themselves, may get poor sleep, and neglect their own well-being (eg, diet and physical activity) [11,14,15]. As such, recent epidemiological research has found spousal caregivers are at increased risk of dementia themselves and has found caregiving to be an independent risk factor for mortality [23,24]. Notably, strategies and interventions are needed to reduce burden and improve outcomes for both the caregiver and care recipient.

Social Media and Caregiver Social Support

Over the past 3 decades, people have increasingly sought health information and support on the web. In high-resource countries, more than half of the adults use the internet for health reasons, often searching for information related to symptomology, diagnoses, and treatment options for health conditions. Online health information seeking is particularly common when users themselves or a close family member have a chronic health condition. While there are numerous websites from which users may access health information, they are increasingly likely to do so specifically via social media platforms.

Defining and delineating what constitutes a “social media platform” has proved challenging; however, most definitions broadly include a limitation to digital technologies, content generation by users, and the capacity for users to interact or share directly with others [25-27]. Over time, the use of social media has exploded, with a vast majority of adults in the United States reporting regular engagement with at least 1 platform and an average daily use of >2 hours [28,29]. Important to the
current research, social media is widely used as a tool for people to connect with mutual friends, interests, circumstances, and hobbies [30-32].

To date, research related to caregivers’ use of social media has spanned many chronic conditions (eg, cancer, diabetes, and physical and mental disability); ages (ie, older adults and children); and platforms (eg, Twitter, X, Facebook, Instagram, and Reddit). Much of this exploratory work has examined use patterns and user needs [33,34], broadly suggesting caregivers most often use social media to exchange information related to care recipient health, psychosocial issues, and daily care activities [33-35]. Among the latter, caregivers may use social media to discuss activities of living, sleep, diet, finance, showering or bathing, transportation, medical care, and formal disease diagnoses [35]. Some research suggests online support is associated with a positive impact on the emotional well-being of medical caregivers, especially for those who used online support for a long term [36]. However, to date, limited literature exists examining the use of social media for support-seeking among AD/ADRD caregivers, specifically [35,36].

AD/ADRD caregivers may be particularly likely to seek support on the web due to myriad factors that limit social support in their offline lives and community settings. For example, AD/ADRD caregivers face documented stigmatization and progressive social isolation as care recipient symptomology worsens and it becomes more challenging to engage in public social settings [6,34,37]. As such, early social interventions for AD/ADRD caregivers have often focused on creating shared spaces with others in similar situations [5,38]. However, rural residents and others who are not located near needed resources (eg, respite or adult day care and support groups) may face challenges in finding others who have similar experiences or a shared understanding of care responsibilities [39]. Furthermore, AD/ADRD caregiving requires substantial time and financial commitments, which create logistical limitations on opportunities for social engagement outside of one’s care responsibilities [23,33,40]. Finally, even in shared spaces with similar others, it may be challenging to express complex or difficult emotions that are common among caregivers [15,22]. Notably, AD/ADRD caregivers may be particularly likely to pursue support on the web via social media, rather than through in-person or community-based programs.

Therefore, this study aimed to explore how caregivers use the social media platform Reddit, with particular emphasis on social support and AD/ADRD information seeking. To this end, we first collected and applied computational natural language processing (NLP) tools to a large corpus of Reddit posts to identify salient themes across the site. We then conducted a manual qualitative content analysis of a random subset (ie, 657/6563, 10.05%) of the corpus to gain a more nuanced understanding of support and information-seeking behaviors of AD/ADRD caregivers through the site.

Methods

Ethical Considerations
For this project, we analyzed only existing data (ie, publicly available information), posted directly to Reddit. All study procedures were approved by the Indiana University Institutional Review Board (#23662).

Data Collection
We collected data for this study from Reddit, a social news aggregation, content rating, and discussion social media platform. Unlike other social media platforms, including Instagram, TikTok, and X (formerly Twitter), Reddit represents a forum of communities where people can “opt in” to subreddits, which are moderated subspaces within the wider platform, dedicated to a certain topic. This allows users to seek community and interaction based on specific interests or needs [41]. Subreddit content varies extensively. However, in recent years, subreddits specific to health and well-being–related topics have become increasingly used by people in need of social support, connection, or solidarity [42,43]. Reddit data are also unique in that posts are unincumbered by length limitations. Therefore, data collected from Reddit tend to be more comprehensive and nuanced compared with data from other social media platforms.

Using the Python Reddit API Wrapper [44], a third-party Reddit data scraper, we identified 2 subreddits specific to our research purpose (r/Alzheimers and r/dementia). We then programmed the Python Reddit API Wrapper to collect new, popular, and trending posts, along with relevant metadata, in both subreddits between May and June 2022. These data were saved as a single CSV file for further computational and qualitative data analysis. Upon completion of data collection and excluding duplicates, blank entries, and those deemed irrelevant to discussions about AD/ADRD, we retained 6593 posts.

Analysis
Given our exploratory research purpose, we undertook a two-step process to analyze our data: (1) computational analyses and (2) manual qualitative coding and review. This 2-step approach afforded the ability to identify topics embedded across the entirety of our data using NLP methods and to understand the meaning of each topic at a deeper and more nuanced level through traditional qualitative analyses (Figure 1). We offer a brief explanation of each method applied to the data in subsequent sections.
Natural Language Processing

**BERTopic**

Topic models refer to any series of NLP tasks that consolidate large language data sets into representative topics or themes. While there are a variety of topic modeling tools, including the highly established latent Dirichlet allocation (LDA) approach, our study used a neural network pipeline leveraging Bidirectional Encoder Representations from Transformers (BERT) vectors using the BERTopic topic modeling tool [45,46]. Indeed, BERTopic is unique among established LDA approaches in that this pipeline specifically uses neural networks and BERT embeddings to approximate latent topics. Topics derived using the BERTopic tool tend to be clearer and more nuanced due to the encoding process, where raw text data are compared against a large language data set. At this stage, raw text data are converted into numerical form, which creates dense vector representations for each document in a given data set. These representations capture contextual nuances of each word across documents, which is not attainable using LDA or other probabilistic topic modeling approaches. Vectors are calculated for each document. As these vectors are difficult to model and interpret, we then perform a dimensionality reduction technique, principal components analysis. Following dimensionality reduction, we extract a series of topics that represent a synopsis of the entire corpus.

**Coherence Score Check**

BERTopic, and most other topic modeling tools, can generate any number of latent topics specified by a predetermined parameter \( k \). However, this fixed topic number may not always reflect the optimal number of topics for a given corpus. Therefore, multiple models may be run and compared, retaining the best overall topic solution for the research. Coherence scores, which are a statistical value predicting the overall interpretability of topics, are a common metric for assessing model fit and choosing the optimal solution. Coherence scores are calculated using pointwise mutual information, which is a measure of the strength of association between 2 words in each document and a collection of documents. In brief, this calculation tells us the degree to which 2 words are more likely to appear together than what would be expected by chance. Therefore, higher coherence scores generally indicate a more interpretable topic solution, as words that co-occur in the corpus more often than they would by random chance are sorted together as topics.

Qualitative Content Analysis

To support findings derived from computational analyses and to better understand nuance related to social support and AD/ADRD information seeking within our data, we conducted a manual qualitative content analysis of a subset of posts previously analyzed. For this analysis, we adopted an approach that was both inductive and deductive in nature. Given our specific research interests related to the ways users sought support on the web, we deductively mapped data codes onto existing conceptual frameworks related to social support and known AD/ADRD behavioral symptomology. However, we inductively allowed for additional codes to emerge from the data.

**Procedure**

**Natural Language Processing**

Once we completed data collection, we began preprocessing our data, a common procedure in computational analyses that seeks to standardize and add cohesive structure to messy text data. Part of this standardization includes removing parts of speech that may detract from the clarity of the models, including first-person names, numbers, special characters, emojis, articles, and a series of stop words. Once our data were preprocessed, we proceeded to analyze the data with the BERTopic tool. To identify the optimal number of topics, we performed an iterative topic model analysis that tests a range of topics by iterations of \( 5 \) (eg, 5, 10, 15, \( \ldots \), \( k \)). For each iteration, we calculated coherence scores. We selected the optimal number of topics based on the highest coherence score among topics, ranging from 5 to 80. Once we identified the optimal topic number, we performed a sorting function, which triaged all tokenized sentences into one of \( k \) latent topics. This sorting feature allowed us to examine topic numbers, keywords associated with each topic, and the number of parsed sentences sorted in each topic. Once computational analyses were complete, we performed a manual review of a random 10.05% (657/6563) of posts within our larger sample. These posts were manually coded and independently compared to the BERTopic output to ascertain overlap.

**Qualitative Content Analysis**

Initial coding was undertaken by 3 individuals on the research team. First, coders trained together on sample data that were not included in later analyses. This initial coding was subsequently reviewed by 2 established researchers who were not initial coders. After review, the research team met 3 times

---

Figure 1. Workflow depicting computational and qualitative data analyses pipeline. PRAW: Python Reddit API Wrapper.
to review the coding of the training data, which helped crystallize code definitions and identify early emergent themes. Any disagreements in the coding of these training data were discussed until a consensus was reached.

After training, a similar coding process was adopted for the larger data set. Initial coding was completed by the same 3 individuals who provided initial open codes in the training data. This initial coding was then reviewed independently by 2 established researchers from the team for consistency and accuracy. We began to draw connections between initial open codes via axial and selective coding processes to arrive at the final structure, as outlined in the Results section. The coding team met twice to discuss the overall coding structure, again resolving disagreements via discussion until a consensus was reached.

Results

Unsupervised Topic Modeling

Our first set of analyses used NLP to identify the broad themes discussed by users in subreddits related to AD/ADRD. The optimal solution, as determined by the coherence score, included 10 topics, with a coherence score of 0.38. Table 1 provides a summary of each topic, including each topic name, keywords associated with each topic, and its proportional representation within the larger corpus. Importantly, our analysis tokenized data at the sentence level. We chose to analyze at the sentence level, rather than the post level, to identify more granular topics within the corpus. Had we analyzed at the post level, the algorithm would have sought to identify the primary overarching theme across the entire post. Because Reddit posts have no limitations on length, users may include multiple topics in a single entry; such intra-post thematic variability would be poorly reflected with a post-level analysis, likely resulting in very general and hard-to-interpret topics. Therefore, to avoid such a loss of nuance, we chose the more granular sentence-level approach.

Proportionally, the topic with the greatest representation in our corpus was topic 0, “Reddit inquiries, narratives, and advice,” comprising 17.45% (16,152/92,562) of sentences in the total corpus. Other prominent topics include topic 1, “venting/expressing feelings and emotions” (12,652/92,562, 13.67%); topic 2, “moving- and housing-related adjustments” (12,090/92,562, 13.06%); and topic 3, “ADRD diagnosis and other acute illnesses” (10,695/92,562, 11.55%). The remaining topics were marginally consistent in terms of distribution, with topic 9, “hospital, assisted living, and memory care,” having the lowest representation of 5.4% (4994/92,562).

Topics derived using any topic modeling tool are potentially overly generalized or difficult to parse if topics share similar keywords. As such, additional review of the topic solution by researchers with topical knowledge is needed to ascribe meaning to each latent topic. To do so, we applied a sorting function based on keywords to sort tokenized sentences into 1 of 10 possible topics. We then reviewed individual posts to understand and contextualize each topic. Table 2 provides excerpts of posts sorted into each topic to assist the reader in understanding the latent topics identified in the corpus. We observed some generalized themes, as well as more specific AD/ADRD caregiving and contextual topics.

The first 2 identified topics are more general in nature. For example, topic 0, “Reddit inquiries, narratives, and advice,” is quite general and reflects the common language of Redditors asking questions to the community. What follows in any given sentence varies (eg, diagnoses, caregiving advice, or difficulties associated with AD/ADRD symptoms), but there was a common theme of seeking, whether related to information or emotional support. Similarly, topic 1, the second most common in the corpus, encapsulated users’ many and varied expressions of emotion. These more generalized themes are often found in the first few sentences of a Reddit post, wherein the user expresses a wider sentiment before providing more specific and varied details.

Another grouping of topics related to the varied contexts of AD/ADRD caregiving and related decision-making. Topic 2, for example, focused on housing- and moving-related concerns, where users often sought information and advice related to the appropriate time to limit a loved one’s capacity to live alone or strategies for ensuring the safety of community-dwelling individuals living with AD/ADRD. Topics 7 and 9, respectively, centered on more professional care settings. Topic 9 focused on hospitals, assisted living, memory care, or other such facilities; many of these posts were regarding providing or seeking information about proper care, identifying a quality facility, financing options, or the appropriate timing for seeking professional care support. Finally, topic 7 included posts related to hospice (ie, end-of-life) care. Notably, users often sought guidance on the appropriate time to move their loved ones to this type of care and emotional support at this challenging stage.

The remaining topics related to different aspects of the caregiving experience, including the impact of diagnosis (and comorbidities) and care on the family, caregiver burden, and 2 general caregiving topics. In topic 3, for example, we observed many posts that related to physical symptoms related to both dementia and related health challenges (eg, injuries from falls). For many who may not have opportunities to connect with similar others, platforms such as Reddit may be the only source of support available (eg, “I really just don’t know where else to turn”). Others either sought or provided specific advice related to managing certain symptoms (eg, aggression and anxiety) and strategies for managing the burdens of care.

Notably, topics 6 and 4 were highly interrelated and discussed a broad set of caregiving-related topics. While the content in these 2 topics was highly related and posts were often similar, topic 4 focused more specifically on women, while topic 6 was less-focused on gender overall. Because of the unsupervised nature of NLP, wherein latent topics are generated based on word co-occurrence but are not constrained to be perfectly orthogonal, it is possible to generate highly correlated topics, as in our data. To further contextualize this overlap (and other bivariate relationships between topics), we generated a correlation matrix using the BERTopic Python module (Figure 2). This figure offers a visual representation of relative topic overlap, with darker colors indicating higher levels of overlap. Not
surprisingly, we can see high levels of overlap between topics 6 and 4. The presence of overlapping or highly correlated topics may imply the existence of more generalized themes or a hierarchical structure, which is beyond the scope of this analysis.

Table 1. BERTopic 10-topic solution, including topic name, top associated words, and overall corpus representation (N=92,562).

<table>
<thead>
<tr>
<th>Topic ID</th>
<th>Topic name / Short name</th>
<th>Top associated words</th>
<th>Sentences, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Reddit inquiries, narratives, and advice / Inquiries</td>
<td>ago, today, night, say, asked, ask, day, asking, guess, idea</td>
<td>16,152 (17.45)</td>
</tr>
<tr>
<td>1</td>
<td>Venting or expressing feelings and emotions / Expressing emotion</td>
<td>situation, feel, say, talk, feeling, felt, saying, talking, thinking, think</td>
<td>12,652 (13.67)</td>
</tr>
<tr>
<td>2</td>
<td>Moving- and housing-related adjustments / Housing or moving</td>
<td>moved, house, moving, apartment, bedroom, phone, home, stay, job, contact</td>
<td>12,090 (13.06)</td>
</tr>
<tr>
<td>3</td>
<td>AD/ADRD(^a) diagnosis and other acute illnesses / AD/ADRD diagnosis</td>
<td>treatment, patient, med, dr, doctor, hospital, diagnosis, medical, appointment, medicine</td>
<td>10,695 (11.55)</td>
</tr>
<tr>
<td>4</td>
<td>Women and caregiving / Women and caregiving</td>
<td>mom, mother, mum, parent, shes, upset, daughter, aunt, angry, grandma</td>
<td>9065 (9.79)</td>
</tr>
<tr>
<td>5</td>
<td>AD/ADRD caregiving burden / Caregiver burden</td>
<td>depression, depressed, cope, stress, stressed, caring, grief, feel, burden, family</td>
<td>7637 (8.25)</td>
</tr>
<tr>
<td>6</td>
<td>Family caregiving / Family caregiving</td>
<td>grandparent, grandchild, sibling, parent, caregiving, grandmother, grandpa, grandma, family, aunt</td>
<td>7624 (8.24)</td>
</tr>
<tr>
<td>7</td>
<td>Hospice care and decision-making / Hospice</td>
<td>caregiving, caregiver, hospice, wheelchair, hospital, aide, health care, rehab, nursing, home</td>
<td>6280 (6.78)</td>
</tr>
<tr>
<td>8</td>
<td>AD/ADRD impact / AD/ADRD impact</td>
<td>dementia, alzheimer, alzheimers, elderly, diagnosed, grandma, grandmother, diagnosis, grandpa, impairment</td>
<td>5373 (5.80)</td>
</tr>
<tr>
<td>9</td>
<td>Hospital, assisted living, and memory care / Professional care settings</td>
<td>Hospital, hospice, grandmother, nurse, nursing, grandma, illness, diagnosis, caregiver, medical</td>
<td>4994 (5.40)</td>
</tr>
</tbody>
</table>

\(^a\)AD/ADRD: Alzheimer disease and Alzheimer disease–related dementias.
### Table 2. Representative post excerpts per topic, derived from our sorting function.

<table>
<thead>
<tr>
<th>Topic ID</th>
<th>Topic name</th>
<th>Representative quotation</th>
</tr>
</thead>
</table>
| 0        | Reddit inquiries, narratives, and advice                 | • “Just curious if anyone has more information.”  
• “I'll probably poke around this subreddit every once in a while.”  
• “Hey Lovelies - just bouncing in to share a little tip that has made a HUGE difference in our household.”  |
| 1        | Venting or expressing feelings and emotions              | • “I made the mistake of lashing out when he asked why I didn't do his stretches Monday morning, saying 'do you really think I have time in the mornings to do that?'...I feel like he only wants me around to do everything for him.”  
• “Like, I have no one to talk to. No one will understand. I just want to run away and be left alone. I did not ask for this to be my life.”  
• “Even before the pandemic I felt like my soul was being pulled from me. I just wish I could be around ‘normal’ people.”  |
| 2        | Moving- and housing-related adjustments                  | • “We recently bought some cameras and a door sensor and suddenly we know that he’ll go outside at 3 AM for whatever reason. Are there locks that can be connected to a CO2 or smoke alarm?”  
• “He also shows no interest in leaving his home; he spends his days in his recliner reading or watching TV.”  
• “I want her to go to AL, which in our area averages about $5k a month [for care]. She has enough money in the bank to live in AL for 4-5 years. She doesn't want to go to AL because ‘It’s too expensive.’”  |
| 3        | AD/ADRD diagnosis and other acute illnesses              | • “They told us they could force-feed her via intravenously or put a tube down her throat or stomach. The doctor’s other suggestion was to ‘make her comfortable.’”  
• “And then one morning...she fell and broke her hip because she didn’t use her walker while heading to the bathroom.”  
• “Symptoms started last November. By January she was hospitalized. After the diagnosis of LBD, medical practitioners started treating her differently.”  |
| 4        | Women and caregiving                                    | • “I picked out a room yesterday while and the guilt is getting to me too, my sibling is having a harder time dealing with it. My heart is breaking over this cat and the bond my mom has with him.”  
• “We gave our youngest kids OUR house and animal because I had made a promise to my mother she wouldn't go into a home years before this s**t got real. My mother was still in total denial, and we were basically ignorant in a not stupid way.”  
• “My grandma is 75 years old and has been acting kinda not herself most of the time since February. But by this time I’ve noticed that she’s often confused, asks the same stuff a lot sometimes, has spontaneous bad mood, crying thinking of something bad.”  |
| 5        | AD/ADRD caregiving burden                                | • “My thoughts run to upping my anxiety meds, cry in my mothers arms, post vague rants on Facebook, depress friends with the tearful story about my situation or pop a Xanax and sleep.”  
• “I always talk about the little adjustments you have to make as your loved one gets worse. I’ve cried for about 50 mins and done the stupid rocking myself back and forth in efforts to comfort myself.”  
• “I feel so powerless and I just want to make you healthy again. Every day I am so happy and put a smile on my face, trying to ignore the gnawing thoughts of your illness and how it is slowly destroying you from the inside.”  |
| 6        | Family caregiving                                        | • “My only experience with this sort of thing is with my spouses family situation. With their family, they do not let the grandma go to her old house any more. I was told by my wife not to bring my mom back to her old home, as it would cause trouble or upset her like how her grandma got upset.”  
• “My father-in-law was hiding from us just how bad it was (Not sure why as we where/are very active in their lives) and until FiL death she put on more of a front for us (didn’t always manage it, but managed enough to hide just how bad it was).”  
• “My dad started answering the phone as my mom became more confused and often asked for her mom (died 50 years ago). Constant desire to go home and find her mom. Not recognizing my dad has been her caregiver through it all.”  |
| 7        | Hospice care and decision-making                         | • “She was released from the hospital into the care of one of her brothers, but as of today she has returned back home on her own. Basically, Family Care is not an option.”  
• “He’s at the stage where he can opt for hospice care, but he’s choosing to prolong life treatment (or at least try.) I told him if he wants to do whatever he wants then hospice care is an option but he will live a lot less longer.”  
• “After days of back-to-back seizures, she is in the hospital completely unresponsive and I got the news they don’t expect she will recover. I have an appointment tomorrow to go over our hospice care options. I thought we had more time. I wont ever hear her voice again. I used to get annoyed at all of the voicemails she would leave because I always called her back as soon as I could.”  |
Representative quotation

• “The husband has dementia and his wife was taking care of him. I could feel the heart break behind her words as her husband had clearly deteriorated over the years to resemble less and less of the person she described.”
• “She passed last night, a decade since it first became clear that the disease that had taken her sister had come for her as well. For most deaths ‘I’m sorry’ suffices, but we all know that with dementia that’s not exactly true.”
• “Over the past four years, my family has come to suspect that my mother has dementia, which doesn’t make a whole lot of sense, given that she’s 52. The further irony is that my grandmother has just now started displaying symptoms of dementia, at almost 90 years old.”

9 Hospital, assisted living, or memory care
• “With my dad in the hospital, my mom has been a complete wreck! The doctor said she’d have to go into a home and she flipped out begging me to bring her to my house.”
• “My mother is in a memory care facility - she’s been there since last July. My father can’t face the fact that this is a behavior situation, not a ‘see yet another doctor for a magic cream that will stop this.’”
• “I have been slowly priming my sister regarding getting mom ready for a home, as she is now truly requiring intense supervision for just about everything and I think my sister is realizing this as her decline has been incredibly apparent over the last few months since dad is gone.”

aAD/ADRD: Alzheimer disease and Alzheimer disease–related dementias.

Figure 2. A 10-topic solution correlation matrix denoting topic similarity.

Qualitative Content Analysis

Overview

Through this analysis, we sought to situate prior NLP findings in the context of existing research on caregiving for people living with AD/ADRD. Specifically, we qualitatively analyzed a subset of our existing corpus, deductively mapping onto existing conceptual models of social support and AD/ADRD behavioral symptomology. We allowed for additional inductive themes to emerge.

Social Support

A total of 2 primary forms of social support were observed among Reddit users: informational and emotional. Of note, companionship behaviors and relatedness (other theoretical forms of social support) were not commonly observed in our data. Among the subcategories of informational forms of social support identified by Yao et al [47], 2 subcategories were
observed frequently among our data: advice and teaching. Users sought advice from others related to a broad variety of topics, including help with decision-making (e.g., seeking professional care support, admission to a care facility, and beginning hospice care); managing challenging symptomology (e.g., aggression and wandering); coordinating care networks; and strategies to prevent burnout. For example, one user posted as follows:

I need help. With anger. With boundaries. Active things I already do: meditation, avoiding most substances, gardening, painting, journaling... tried therapy a few times they always offer meds which I dont [sic] want due to chronic illness I dont [sic] want to flare up... thank you.

Beyond this more generalized advice seeking, users sometimes sought or provided specific forms of teaching informational support. Often, these threads surrounded discussions of new medications or clinical trials for people living with dementia. Others offered practical tips and supports. For example, one user described the placement of an “orientation board” in a conspicuous place to provide useful information for their care recipient, including the day or date, expected appointments, weather, and other basics that helped reduce confusion.

In addition, we observed users both seeking and providing emotional social support. We observed a broad array of emotional support, including words of affirmation, expressions of care or concern, encouragement, sympathy, and understanding. Frequently, users expressed a vague or nonspecific need for support—that they had come to the forum seeking others with similar experiences who understood the challenges associated with caregiving. However, many users noted the sense of community and the value of the group’s shared experience in helping them individually. For example, a user thanked the group, saying the following:

I appreciate everyone so much. I know that what were [sic] experiencing with loved ones and or friends is painful, tiring, and other things. Your questions and experiences have given me a heads up on situations.

In a different case, users specifically coalesced and supported a member who was going through a particularly challenging period, including providing donations and financial support. The user thanked the group, saying the following:

Feeling a little better today and am very overwhelmed by the support. You all really did make me feel less alone and I appreciate your support and offers to talk when you could have easily scrolled past. I also appreciate the awards but next time, take that money and please donate it to the Alz [sic] association. I do love the hug and silver award though, that is very very kind.

AD/ADRD Behavioral Symptomology

Overview

We coded AD/ADRD behavioral symptoms using the standard Neuropsychiatric Inventory [48]. We found users regularly discussed symptomology to approximate or assess disease progression, particularly in the context of major decision-making (e.g., care levels and power of attorney). A total of 3 categories of neuropsychiatric behavioral symptoms were commonly observed in Reddit posts.

Agitation or Aggression

One primary behavioral concern that appeared consistently in the discussion was increasing levels of aggression or agitation. Users noted that as part of the disease progression, care recipients increasingly struggled with emotional regulation; this often led to outbursts and physical aggression. Users recalled incidents where care recipients had yelled at caregivers, thrown objects, and attempted physical violence—sometimes without explanation. For example, one user described attempts to keep a loved one from wandering by securing doors and fences on the property; during the resulting “tirade,” the care recipient “grabbed my brother by the front of his shirt and began screaming in his face” and they also noted the following:

He has threatened to slash the tires on the car. The other day I caught him just before he swung a metal pipe at the windshield of my rental car.

Among Reddit users, we found these discussions were often related to larger caregiving decisions. For some, the emergence of violence was an inflection point in seeking full-time or professional care. However, others sought help, as the aggressive behaviors of care recipients caused them to be removed from care facilities.

Delusions and Hallucinations

Another prominent theme in the data was increasing delusions among care recipients. These delusions often manifested as care recipients’ mistrust of those around them due to unfounded beliefs that these individuals would harm them in some way (e.g., physical violence and theft) or as a conspiracy against them. Episodes of delusion covered a broad range of experiences, from a more benign suspicion that others were lying to more elaborate, reoccurring fantasies. For example, one user posted regarding regular hallucinations their parent had about a fantasy world:

[A] fantasy world he created involving red-robed people who came into the room uninvited and stayed overnight.

Regarding aggressive behaviors, many discussions around delusions and hallucinations were part of larger threads seeking support and understanding about dementia progression. Delusions were regularly discussed as a motivation for seeking external care support.

Nighttime Behaviors

The final commonly discussed behavioral symptom category related to nighttime issues. In the context of dementia, this category relates to a broad set of behaviors, including rising too early in the morning, awakening during the night, and excessive daytime napping—all generally related to sleep disturbance of some kind. Users regularly discussed the progression of sundowning—wherein individuals may experience restlessness, confusion, or irritability in the evening as daylight fades and progresses into the night. Users noted that sundowning symptoms continued into nighttime and often caused an inability...
to sleep or that patients would experience nightmares or terrors and stay awake throughout the night. In some cases, this inability to sleep resulted in the individual becoming fearful, seeking to escape, or taking additional medication. For example, one user noted the following about their mother:

[K]eeps getting up in the middle of the night, walking unassisted, digging around in her med suitcase taking God knows how many extra meds in the middle of the night. [sic]

Commonly, these nighttime behaviors also had negative health repercussions for caregivers and family members who could not sleep themselves, either due to disturbances caused by the care recipient or by the need to provide supervision and care while the individual was awake.

Emergent Codes
Caregiver Anxiety or Guilt
The most frequently observed code across the entire data set related to caregiver guilt or anxiety, with users expressing distress related to their caring role and responsibilities, as well as their own emotions and personal challenges. For example, some users expressed fear of failing in their caregiving role, such as a user who expressed, “My biggest fear is that I’m not doing enough,” or another who said the following:

I’m struggling hard tonight. I want to want to fix this for my dad so badly but I can’t and I hate it. I hate that he has no idea what’s going on or why its happening and I hate that I can’t just make it all better for him. [sic]

Another primary form of guilt and anxiety that emerged related to the users’ emotions and feelings. Some felt anger toward their care recipient. For example, some users noted they experienced emotional exhaustion related to caregiving and challenges coping with the loss of a loved one. One user, for example, lamented the loss of their mother’s cognitive function and was struggling with the decision to place her in an assisted living facility. The user experienced guilt around their feelings, noting the following:

To me, my mother has been gone a long time. Am I an asshole for not wanting to see her...? [sic]

Furthermore, multiple users expressed guilt associated with feelings of relief (both realized and expected) upon the passing of the person in their care, such as the user who stated the following:

I feel sad but also I feel free...I think sometimes “now I can start my life” and then I feel selfish and sad.

Decisions-Making About Care Facilities
Another commonly observed theme in the data involved decision-making about care facilities and other forms of assisted living. Prospectively, users sought advice and validation related to deciding when in the disease progression it was appropriate to place their loved ones in an assisted living facility. In addition, users sought practical advice on finding and selecting the right place and evaluating quality of care. However, many posts discussed challenges related to patients already in an assisted living facility. For some, there were unexpected or short-notice care expectations, even after the individual was placed in an assisted living facility. One user noted as follows:

The facility staff are telling me we either need to coordinate a sitter to literally sit outside his door 24/7 and physically keep him from leaving, or to send him to a mental institution until they get his meds right. This change needs to happen within 24 hours.

Other users were disappointed with the quality of care their loved ones were receiving. For example, one user said the following:

She was able to get out via an alarmed door, which apparently didn’t sound, get past reception, out the front door and walked up a super busy street to a grocery store 1/2 mile away! The street is almost a highway with 3 lanes of traffic on each side. I don’t know all of the details yet but clearly this is not acceptable...I’m so angry that this could happen. My family is paying a ton of money to keep them safe and this happens.

Therefore, the myriad challenges associated with care facilities were common themes in the discussion on dementia-related subreddits.

Legal and Financial Planning
Users also commonly discussed the challenges associated with legal and financial planning for their loved ones living with AD/ADRD. These posts largely centered on the high financial costs of professional care and assisted living facilities and the steps caregivers and families could take to afford this support and protect family assets. Many users noted that their loved ones did not have sufficient resources (eg, retirement savings, long-term care insurance, and state support) to cover the costs of needed full-time care. In some cases, users reported subsidizing their loved ones’ care with their own funds:

I used to be what I would consider reasonably well off, now I have nothing left and struggle daily to make ends meet.

Others described an impossible tension, wherein they could neither afford to pay for needed care nor quit their own jobs to provide full-time care, such as the user who said the following:

My only word of advice is, unless you’re a millionaire and can afford home care, which we were most certainly not, there is no solution and no easy road to this disease.

One common concern in the discussion of care was the protection or leveraged use of assets, specifically houses owned by the care recipient. For example, one user described a situation in which their parent could not live in his home because he needed rental income to pay for his care, effectively forcing him to move in with relatives. Others noted that owning a home precluded access to social support programs as it counted as a substantial asset that had to be depleted before receiving assistance. Several threads related to strategies for protecting such assets (eg, placing the home in a trust) or challenges in liquidating them (eg, relatives living in the house or

https://aging.jmir.org/2024/1/e59294
unwillingness to sell). Furthermore, these conversations commonly discussed the need to get legal affairs in order, such as establishing power of attorney and writing a living will early in the disease progression.

**Thematic Comparison**

Across both forms of data analysis, we observed similar overall topical structures within Reddit posts related to AD/ADRD caregiving. The most salient NLP-identified topics were very general and focused on making inquiries or sharing emotions with the forum. In our qualitative analysis, the most common coding was quite similar, related to users’ expressions of emotion. However, through qualitative coding, we identified added nuance to these emotions of guilt, anxiety, and the struggles of AD/ADRD caregiving. Furthermore, we noted users sometimes expressed feelings that may be socially unacceptable outside an anonymous online space. For example, some users expressed relief upon the death of their care recipient. While likely not an uncommon emotional response given the myriad challenges associated with care, this relief may be uncomfortable to express to others for fear of judgment.

Similarly, the corpus-wide NLP analysis identified information seeking as a primary use of Reddit forums for AD/ADRD caregivers. However, in our qualitative coding, we identified specific informational needs of caregivers (eg, legal and financial planning advice) that were not apparent in the wider analysis. Moreover, we found through the qualitative analysis that users often sought normative standards for decision-making at common key time points in the disease progression, such as the “right” time to move a care recipient to professional or end-of-life care. While we observed themes related to assisted living and hospice care in the NLP, the qualitative coding provided a deeper understanding, finding that users were most often seeking external standards from others, which they could then apply to their unique situations and reduce doubt in their own decision-making.

Therefore, our broad NLP analyses gave a high-level, atheoretical overview of themes in a large corpus beyond what is generally feasible for manual, human-driven analysis. Our qualitative coding fleshed out these findings, incorporating existing conceptual frameworks to contextualize and provide a more detailed understanding of the social support and information-seeking behaviors of AD/ADRD caregivers on Reddit.

**Discussion**

**Principal Findings**

Users sought community and informational support in dealing with the challenges associated with caregiving for people living with AD/ADRD. Consistent with prior literature, users reported a variety of emotional challenges related to providing care [34-36]; caregiver guilt and anxiety was a highly salient theme across the forum. The sources and manifestations of these emotions were highly variable across the forum. For example, users noted feelings of inadequacy in their capacity to care for loved ones, guilt associated with a lack of patience for their care recipient, loneliness and social isolation, burnout, and anxiety related to the varied stressors associated with care. While caregiver burden is well-established in the literature [15], it was particularly salient in the online space, wherein this seemed to be a primary point of conversation between users. Given the documented social stigmatization of dementia caregivers [37], users may seek to express these feelings and find support online as they do not feel comfortable expressing their challenges in day-to-day life. Furthermore, and unique to the anonymous online space, users felt empowered to discuss potentially taboo topics, including relief upon the passing of their care recipient.

It is important to note the subreddits analyzed in this study are moderated to avoid bullying and spam and to establish a respectful dialogue in the forum. Notably, users may feel particularly safe to express these otherwise stigmatized beliefs and challenges. Creating this trust and safe space to discuss the complex emotions associated with dementia care may be similarly useful in face-to-face caregiver support interventions.

Consistent with other studies, Reddit users discussed myriad behavioral symptoms of AD/ADRD and strategies for providing effective care related to each symptom [19,20]. Manual coding identified instances of 14 different behavioral symptoms of dementia. The most identified among these were agitation or aggression, delusions, and bathroom or toileting issues. These are, perhaps, not surprising, given that these are symptoms with higher associated costs or challenges of care. Therefore, these may be the symptoms most discussed when seeking support. Often, users reported multiple symptoms together in the same post, consistent with the nonlinear nature of disease progression. Many users sought to map symptomology onto a disease progression timeline to better understand their own situation. However, as was noted by multiple users, this is challenging as dementia can present and progress differently across individuals.

As such, the presentation of behavioral symptomology was often highly variable across the sample.

Finally, discussions with the forum also often related to difficult decisions facing caregivers of people living with dementia [4,49]. Many posts focused, for example, on the legal and financial planning challenges of care; users provided resources to others to help create power of attorney documentation and to navigate health care systems. The complicated legal and financial challenges associated with dementia and end-of-life care have been documented elsewhere. In our sample, users sought to crowdsource resources to navigate these complex systems, potentially without the high costs associated with hiring specialist attorneys and personal financial planners. Furthermore, users sought feedback associated with care decisions, particularly the proper timing for moving their loved ones into a full-time care facility. User discussions often compared symptomology and family circumstances, seeking validation or support in making the choice to seek professional care help. Again, these discussions often centered on stigma, fears associated with care quality, and the financial costs of professional care. As users faced difficult but common decisions related to transitions in care—times associated with higher perceived caregiver burden—they sought online emotional and informational support.

**Implications for Practice**

Our study findings suggest the need for increased social support structures for AD/ADRD caregivers, a special population who

https://aging.jmir.org/2024/1/e59294
may experience challenges associated with seeking information and other forms of support on the web. This is consistent with the wider literature surrounding caregiver burden, suggesting high rates of burnout and needs for additional support beyond localized networks [34,35]. However, our study findings introduce important additional considerations. Recent literature has begun to document mental health risks associated with extensive social media use, particularly linked to unfavorable social comparison (ie, to others with fewer perceived challenges) [50]. There are further risks associated with seeking health-related information on the internet, given the high rates of misinformation found on social media [51]. Therefore, despite the somewhat closed and moderated nature of dementia-related subreddit forums, the overall absolute value for AD/ADRD caregivers of seeking support and information therein is unknown. Online resources and spaces for social connection, such as Reddit, which also feature expert fact-checking and seek to reduce harmful social comparisons, may be needed to support caregivers. Furthermore, there are opportunities for the development of mobile apps specific to AD/ADRD caregiving; however, additional research on the relative benefit of new mobile apps versus commonly used social media platforms is needed [40].

In addition, users noted specific informational needs with which they struggled to find high-quality support, including legal or financial planning and strategies for coping with feelings of inadequacy or guilt. These gaps in readily available information offer opportunities for intervention. For example, programs designed to teach caregivers about complex legal and financial planning documents (eg, power of attorney, living wills, and advance directives) may reduce burden. Given the well-documented financial strain associated with providing care, a centralized, online information hub for legal and financial planning may increase the capacity to proactively manage the complex tasks associated with estate planning and long-term care financing. Furthermore, specialized support groups for caregivers may use acceptance and commitment therapy techniques, which emphasize self-compassion around perceived personal shortcomings [52]. Acceptance and commitment therapy techniques are increasingly used in family caregiving settings (including for dementia), with promising early acceptance or feasibility and results related to psychological flexibility, which is the capacity to stay in contact with the present, irrespective of negative thoughts or feelings [52-54]. Unfortunately, access to such interventions and therapies remains limited, leading caregivers to seek information and social support in free and open online spaces. Future intervention work may leverage such opportunities to create readily accessible supports to assist caregivers in managing both specific care tasks (eg, financial management) and emotional challenges associated with caring for a loved one living with AD/ADRDP2.

Limitations and Future Directions
As with all research, this study has certain limitations that should be considered. First, the sample was drawn from a single, moderated, online space, which may limit some of the most extreme responses from being published. Reddit users are, on average younger, highly educated, and more likely to be male than the general population. Furthermore, because data were published anonymously and are self-reported, we have no mechanism by which to directly assess truthfulness. However, this anonymity was a key feature to our findings and has been reported as a strength in other social media studies. Future research may explore AD/ADRDP2 caregivers’ use of different social media platforms, as social media use varies across demographic groups. Similarly, future research may examine differences between social media platforms, due to the differing nature of the content (ie, images, text, and video) found on each platform. In addition, while we analyzed a large sample using computational methods, the qualitative process for deeper analysis was limited to a smaller subsample. Therefore, our findings are not necessarily representative of all dementia caregivers and should be considered with these limitations in mind.

Conclusions
In this study, we used both computational and traditional qualitative analyses to explore the experiences of caregivers of people living with AD/ADRDP2 who posted to the social media channel Reddit. Using an unsupervised topic modeling approach, we generated a 10-topic solution (coherence score=0.38) from a corpus of more than 6500 posts. These topics broadly centered on emotional and logistical challenges associated with care, as well as a cluster of topics associated with various symptomology. To gain a deeper understanding of these topics, we conducted a qualitative review of a subset of posts. Users noted high levels of burden, guilt, and anxiety associated with caregiving. Users sought emotional and informational support to manage the behavioral symptomology of their care recipients and to make key legal, financial, and other care-related decisions. Interestingly, the anonymous and moderated nature of the Reddit platform seemed to reduce the perceived risk of stigmatization, allowing users to express difficult and complex emotions related to the care experience, including resentment and relief upon the passing of a loved one. These findings suggest the need for additional caregiver support interventions to reduce burden and improve overall well-being for both the caregiver and care recipient.

Acknowledgments
ACP is supported, in part, by the National Institute on Aging of the National Institutes of Health under award number K01AG084930. This work is the responsibility of the authors and does not represent the official views of the National Institutes of Health or the National Institute on Aging.


42. Valdez D, Patterson MS. Computational analyses identify addiction help-seeking behaviors on the social networking website Reddit: insights into online social interactions and addiction support communities. PLOS Digit Health 2022 Nov 9;1(11):e000143 [FREE Full text] [doi: 10.1371/journal.pdgh.000143] [Medline: 36812569]


Abbreviations

AD/ADRD: Alzheimer disease and Alzheimer disease–related dementias
BERT: Bidirectional Encoder Representations from Transformers
LDA: latent Dirichlet allocation
NLP: natural language processing
Internet Use as a Moderator of the Relationship Between Personal Resources and Stress in Older Adults: Cross-Sectional Study

Angélique Roquet¹, PhD; Paolo Martinelli¹, MA; Charikleia Lampraki², PhD; Daniela S Jopp¹, PhD

¹Institute of Psychology, University of Lausanne, Lausanne, Switzerland
²Department of Psychology, University of Geneva, Geneva, Switzerland

Corresponding Author:
Angélique Roquet, PhD
Institute of Psychology
University of Lausanne
Rue de la Mouline
Lausanne, 1015
Switzerland
Phone: 41 774602626
Email: angelique.roquet@unil.ch

Abstract

Background: Internet use has dramatically increased worldwide, with over two-thirds of the world’s population using it, including the older adult population. Technical resources such as internet use have been shown to influence psychological processes such as stress positively. Following the Conservation of Resources theory by Hobfoll, stress experience largely depends on individuals’ personal resources and the changes in these resources. While personal resource loss has been shown to lead to stress, we know little regarding the role that technical resources may play on the relationship between personal resources and stress.

Objective: This study aims to investigate the moderating effect of technical resources (internet use) on the relationship between personal resources and stress in younger and older adults.

Methods: A total of 275 younger adults (aged 18 to 30 years) and 224 older adults (aged ≥65 years) indicated their levels of stress; change in personal resources (ie, cognitive, social, and self-efficacy resource loss and gain); and internet use. Variance analyses, multiple regression, and moderation analyses were performed to investigate the correlates of stress.

Results: Results showed that older adults, despite experiencing higher levels of resource loss (questionnaire scores: 1.82 vs 1.54; P<.001) and less resource gain (questionnaire scores: 1.82 vs 2.31; P<.001), were less stressed than younger adults (questionnaire scores: 1.99 vs 2.47; P<.001). We observed that the relationship among resource loss, resource gain, and stress in older adults was moderated by their level of internet use (β=.09; P=.05). Specifically, older adults who used the internet more frequently were less stressed when they experienced high levels of both loss and gain compared to their counterparts who used internet the less in the same conditions. Furthermore, older adults with low resource gain and high resource loss expressed less stress when they used the internet more often compared to those with low internet use.

Conclusions: These findings highlight the importance of internet use in mitigating stress among older adults experiencing resource loss and gain, emphasizing the potential of digital interventions to promote mental health in this population.

(JMIR Aging 2024;7:e52555) doi:10.2196/52555

KEYWORDS
internet use; aging; stress; personal resources; technical resources; Hobfoll’s Conservation of Resources theory; COR theory

Introduction

Background

Internet use has increased drastically in the last decade, with two-thirds of the world’s population now being familiar with it [1]. For example, in Switzerland, 90% of people aged ≥15 years use the internet daily, and even 70.3% of people aged ≥65 years use it regularly [2]. However, few studies have explored the role of internet use on psychological processes such as stress in the context of personal resource loss and its age-related specificities.

Older adults are often confronted with a variety of challenges that can result in social, health, and cognitive losses. Specifically, older adults may experience reduced social support...
networks, decreased physical functioning, and a decline in cognitive abilities such as memory and executive functioning [3,4]. These losses can have a negative impact on overall well-being, including increased levels of stress, depression, and anxiety [5,6]. Moreover, individuals tend to place greater value on avoiding losses than on acquiring gains and, as a result, are often more motivated to take action to avoid potential losses than to pursue potential gains. Technical resources can serve as an additional resource that helps older adults maintain or improve their level of mental health in various ways [7,8], such as through web-based social interaction, web-based counseling, health services, or cognitive stimulation. In other words, technical resources can help older adults reach their maintenance and loss management goals.

In this study, we investigated the effect of technical resources, specifically internet use, and the extent to which they can buffer the relationship between personal resources and stress in both younger and older adults.

**Age-Related Differences in Personal Resources and Their Impact on Stress Levels**

Personal resources are typically considered attributes that individuals value and that enhance their ability to function effectively in terms of controlling and impacting their environment [9,10]. Moreover, individuals' personal resources, including their health, social support, and financial means, along with their mental strengths such as self-efficacy, change over the course of life due to a combination of factors, including biological aging, life experiences, and environmental factors. According to the life span theory, personal resources tend to decrease as individuals age, putting older individuals at risk of decline in overall well-being [11]. For instance, age-related changes such as declining social networks, poor mobility, retirement, and development of chronic illness can contribute to social isolation and feelings of loneliness in older adults [12-14], which have been associated with increased stress and poor well-being [15-17].

Various theories suggest that older adults may develop compensatory strategies to manage the decline in personal resources and maintain well-being [5,18-22]. According to the Conservation of Resources (COR) theory by Hobfoll [20], individuals seek to gain new resources to maintain or enhance their well-being, particularly in the face of stress and adversity [6]. More specifically, when individuals experience a loss of resources, such as a decline in health or social support, they may be more vulnerable to stress and negative well-being outcomes. However, if they can gain new resources, this can help offset the negative impact of the loss and buffer against the effects of stress. Thus, while resource losses can have a significant negative impact on individuals' well-being, resource gains can help replenish those losses and promote resilience.

For older adults, resource gains may involve engaging in activities such as taking classes, volunteering, participating in social activities, or learning new skills. Specifically, technical resources provide a support to access new activities [23]. However, the extent to which internet use influences the relationship between personal resources and stress in aging is not yet fully understood. This study aimed to better understand how the internet can be used as a resource to support the well-being of older adults, particularly in the context of stress and aging.

**Internet Use Role as a Resource Gain in the Context of Older Adults' Stress**

Internet use can facilitate the gain of resources, providing older adults with additional means to cope with stress and improve their mental health outcomes [24-29]. For instance, the internet can serve as a platform for social support, information seeking, and engaging in meaningful activities, all of which can contribute to better mental health outcomes in later life [23]. More specifically, higher levels of internet use predict higher levels of social support, reduced loneliness, and better life satisfaction and psychological well-being among older adults [26].

For example, by using the internet for social interactions, older adults can increase their social networks, receive emotional support, and build relationships with others [28-30] to reduce social isolation and stress. For example, Li et al [31] examined the relationship among social isolation, cognitive functioning, depression, and internet use among older adults. The results showed that social isolation was significantly associated with poorer cognitive functioning and higher levels of depression among older adults. Moreover, internet use moderated the relationship between social isolation and cognitive functioning, suggesting that internet use may have a protective effect on cognitive functioning among socially isolated older adults. Finally, the results showed that internet use was associated with lower levels of depression among older adults regardless of their level of social isolation. This suggests that internet use could serve as a protective factor for cognitive functioning and that it represents an important factor for improving mental health outcomes among older adults.

Moreover, through information research using the internet, older adults can gain knowledge and skills to manage stressors that arise in later life. Being confronted with or anticipating age-related loss of physiological functioning, older adults are interested in acquiring health knowledge [32-34]. Higher web-based health literacy is associated with more positive health behaviors and better health knowledge and attitudes in older adults [35]. In addition, engaging in other meaningful activities such as web-based learning, gaming, shopping, and hobbies can also have positive effects on the psychological outcomes of older adults [36-41]. For example, Gallistl and Nimrod [36] examined the relationship between internet use for leisure activities and well-being among older adults. The results showed that older adults who used the internet for leisure activities reported higher levels of subjective well-being, social connectedness, and life satisfaction than those who did not use the internet for these activities. Moreover, several studies have reported that web-based gaming improves older adults' physical and cognitive functioning [42,43], social interaction, and enjoyment and decreases social isolation [44-46]. These studies commonly suggest that, by participating in web-based activities, older adults can gain new skills, challenge themselves, increase their social networks, and find enjoyment in their free time.

https://aging.jmir.org/2024/1/e52555

JMIR Aging 2024 | vol. 7 | e52555 | p.586

JMIR Aging 2024 | vol. 7 | e52555 | p.586
Little research has been conducted on the relationship among personal resources, internet use, and stress at different ages. However, determining the potential benefits of internet use in managing stress in later life may highlight the importance of promoting access to and use of technology among older adults.

This Study
The main objective of this study was to better understand the underlying mechanisms contributing to age-related differences in stress as a function of changes in personal resources and internet use. More specifically, we investigated (1) age-related differences in stress and its associated predictors; (2) whether the level of resource gains buffered the relationship between resource losses and stress in younger and older adults, replicating the findings by Hobfoll [20]; and (3) whether the profile of internet use in younger and older adults moderated the relationship between resource gains and losses and stress levels.

First, we tested the hypothesis that stress levels differed by age group. We expected that older adults would report lower levels of stress compared to younger adults. This hypothesis was based on previous research that has consistently demonstrated that older adults are exposed to fewer stressors than younger adults [46,47], leading to better well-being outcomes such as less stress [48-50].

The second set of hypotheses concerned the replication of the COR theory [20], describing that resource losses would have a considerably stronger impact than resource gains on individuals’ stress perception [6]. Moreover, COR theory explains that resource gains buffer the effect of resource losses on stress [20]. Accordingly, we expected that (1) more resource losses would be associated with a higher level of stress; (2) resources losses would have a stronger impact on stress than resource gains; and (3) the relationship between resource losses and stress would be moderated by the level of resource gains, with higher levels of gains helping buffer the negative impact of losses on stress levels.

Finally, we hypothesized that the moderating effect of resource gains on the relationship between resource losses and stress would vary based on levels of internet use, presenting distinct profiles for the younger and older adults. First, given previous findings showing that internet use may have a positive impact on older adults’ well-being and stress levels [51-53], we expected that internet use would moderate the relationship between resource gains and losses and stress—individuals who reported higher levels of internet use would experience a greater protective effect of resource gains (ie, stronger effect of gains) than those who reported lower levels of internet use. This hypothesis was based on previous studies demonstrating that higher use of the internet was associated with higher levels of stress, depression, loneliness, and anxiety in young adults [54,55].

Methods

Procedures and Participants
We conducted a cross-sectional study in the French-speaking part of Switzerland. The participants were native or fluent French speakers. We recruited 510 individuals, of whom 280 (54.9%) were aged <30 years (mean age 25.00, SD 2.09 y) and 230 (45.1%) were aged >65 years (mean age 73.55, SD 7.16 y; see Table 1 for detailed participant characteristics). Younger participants were mainly undergraduates from the University of Lausanne, whereas the rest of the participants were recruited using the snowball sampling technique [56]. Recruited individuals volunteered to participate in the study and were not remunerated. For being included, participants had to be able to speak and understand French and had to be aged between 18 and 30 years or >65 years. Participants filled out a web-based open questionnaire containing questions on stress and potential predictors or moderators such as personal resources and internet use (ie, questions presented in a specific order).

The sample size used in this experiment was based on an a priori power analysis conducted in G*Power (version 3.1) [57]. We assumed an effect size of Cohen $f=0.06$, which was derived from previous relevant studies on the buffer effect of resource gains on the association between resource losses and stress [58,59], and an $\alpha$ of .05. Specifically, a total sample size of 404 participants (n=202 per group) provided 90% power to detect effects. To exceed this criterion and achieve >80% power, we recruited 510 participants (ie, n=280, 54.9% younger adults and n=230, 45.1% older adults).
Table 1. Participant characteristics (N=510).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Younger adults (n=280)</th>
<th>Older adults (n=230)</th>
<th>Chi-square test for group comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y), mean (SD; range)</td>
<td>25.00 (2.09; 18-29)</td>
<td>73.55 (7.16; 65-98)</td>
<td>Chi-square (df)</td>
</tr>
<tr>
<td>Gender, n (%)</td>
<td></td>
<td></td>
<td>P value</td>
</tr>
<tr>
<td>Women</td>
<td>190 (67.9)</td>
<td>130 (56.5)</td>
<td>.01</td>
</tr>
<tr>
<td>Men</td>
<td>90 (32.1)</td>
<td>96 (41.7)</td>
<td>.03</td>
</tr>
<tr>
<td>Educational levelb, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obligatory school not finished</td>
<td>1 (0.4)</td>
<td>16 (7)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Obligatory school</td>
<td>4 (1.4)</td>
<td>51 (22.2)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Professional formation</td>
<td>18 (6.4)</td>
<td>79 (34.3)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>General education</td>
<td>4 (1.4)</td>
<td>5 (2.2)</td>
<td>.56</td>
</tr>
<tr>
<td>Professional maturity</td>
<td>15 (5.4)</td>
<td>10 (4.3)</td>
<td>.60</td>
</tr>
<tr>
<td>Gymnasium maturity</td>
<td>49 (17.5)</td>
<td>9 (3.9)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Specialized university</td>
<td>15 (5.4)</td>
<td>20 (8.7)</td>
<td>.14</td>
</tr>
<tr>
<td>University</td>
<td>173 (61.8)</td>
<td>37 (16.1)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Doctoral degree</td>
<td>0 (0)</td>
<td>3 (1.3)</td>
<td>.06</td>
</tr>
<tr>
<td>Financial adequacy, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More money than needed</td>
<td>54 (19.3)</td>
<td>59 (25.7)</td>
<td>.08</td>
</tr>
<tr>
<td>Enough money</td>
<td>177 (63.2)</td>
<td>158 (68.7)</td>
<td>.25</td>
</tr>
<tr>
<td>Less money than needed</td>
<td>48 (17.1)</td>
<td>13 (5.7)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

aChi-square analyses were not conducted on participants’ mean ages.

bThe term obligatory school not finished corresponds to <11 years of education; obligatory school corresponds to 11 years of education; professional formation, general education, professional maturity, and gymnasium maturity correspond to 4 additional years of education; university and specialized university correspond to 3 to 5 additional years of education; and doctoral degree corresponds to 3 to 5 additional years of education.

Ethical Considerations

This study was approved by the Social and Political Sciences Ethics Committee of the University of Lausanne (C-SSP-092022-00002). Written informed consent was obtained from participants before the questionnaire was made available, and they were informed that they could decide to quit the study at any point. This ensured that participants were well informed of the study’s objectives and the potential impact of their contribution. Furthermore, participants were made aware of the duration of their involvement, which entailed completing a web-based questionnaire lasting approximately 30 minutes (ie, the questionnaire comprised 27 pages with 15 items per page and the possibility to go back).

Before deployment, the questionnaire was tested to ensure its feasibility, enhancing its clarity and ease of completion. For nonapplicable items or when participants chose not to respond, options such as “not applicable” or “prefer not to say” were provided, respecting participant autonomy while preserving data integrity. In addition, to ensure valid responses, at least one answer selection per question was mandated, minimizing incomplete or inconsistent submissions and maintaining data reliability.

The informed consent process outlined the data management protocols, including the types of data collected, the methodologies used for data treatment using SPSS (IBM Corp), and the storage solutions provided by Switch Drive (Switch). A commitment was made to the participants that their data would be anonymized and held confidentially, with plans for eventual sharing in an open-access data repository (eg, SWISSUbase for 5 years) after the removal of any personally identifiable information. This study was developed using SurveyMonkey (for the Checklist for Reporting Results of Internet E-Surveys, see Multimedia Appendix 1), a web-based survey platform known for its ease of use and robust data analysis tools. It allows for the creation, distribution, and analysis of surveys, making it an ideal choice for collecting detailed feedback and insights. In addition, SurveyMonkey’s strong emphasis on data security and privacy ensures the integrity and confidentiality of the data collected in the study. Multiple submissions were controlled by monitoring IP addresses and the anonymous codes assigned to each participant in addition to checking for consistency in the responses. Moreover, analyses were performed on questionnaires that were fully completed. No monetary compensation was provided to participants upon the completion of the questionnaire.
Measures

Predictors

Sociodemographic Variables

Demographic variables included age (in years), gender (0=men; 1=women), educational level (1=obligatory school not finished, 2=obligatory school, 3=professional formation, 4=general education, 5=professional maturity, 6=gymnasium maturity, 7=specialized university, 8=university, and 9=doctoral degree), and financial adequacy (1=more money than needed, 2=enough money, 3=less money than needed).

Personal Resources

Personal resources were assessed using the 13-item Personal Resource Questionnaire—Short Form [60]. The short version of the questionnaire includes items concerning cognition, self-efficacy, and social relations. For this study, we used 2 parts of the questionnaire: losses (“To what extent did the listed resources decrease in the last year?”) and gains (“To what extent did the listed resources increase in the last year?”); a total of 13 items × 2 = 26 items; for details, see Table 2). Each item was evaluated on a 5-point Likert scale ranging from 1=none to 5=great amount. Mean composite scores were calculated for losses and gains. Specifically, we created 2 types of indicators (ie, domain-general resources=resource gains and losses) and 3 types of resources in each condition (ie, domain-specific resources=self-efficacy, cognition, and social resources). Lower scores indicate lower levels of each personal resource condition.

Table 2. Details of the Personal Resource Questionnaire used in this study.

<table>
<thead>
<tr>
<th>Resource</th>
<th>Number of items</th>
<th>Items</th>
<th>Cronbach α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognition</td>
<td>5</td>
<td>“Sound cognitive functioning”</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Intelligence”</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Good memory ability”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Ability to concentrate”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Ability to think and understand quickly”</td>
<td></td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>4</td>
<td>“Sense of control over my life”</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Ability to control my future”</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Ability to achieve my goals”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Ability to put my plans into action”</td>
<td></td>
</tr>
<tr>
<td>Social relations</td>
<td>4</td>
<td>“Companionship of other people”</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Close relationship to at least one friend”</td>
<td>0.64</td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Positive relationship partner”</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>“Close relationship to one or more family members”</td>
<td></td>
</tr>
</tbody>
</table>

Internet Use

Internet use was assessed using the 8-item Mobile Device Proficiency Questionnaire [61] measuring participants’ ability to perform on the Internet using a mobile device (example item: “Using a mobile device I can read the news on the Internet?”). Each item was scored on a 5-point Likert scale ranging from 1=never tried to 5=very easily. A mean composite score was calculated, with lower scores indicating low levels of internet use. The Cronbach α for this study was 0.96.

Outcome Variable: Stress

Participants’ stress level was assessed using the 5-item Perceived Stress Scale by Cohen et al [62] (eg, “In the last month, how often have you felt anxious and stressed?”). Each item was scored on a 4-point Likert scale ranging from 0=never to 4=very often. We calculated a mean composite score in which lower values indicated a less frequent experience of stress in the last month. The Cronbach α was 0.80.

Analytical Strategy

Differences between younger and older adults were first tested on stress and its predictors or moderators (eg, personal resources) using between-group ANOVAs. We then conducted correlation analyses to gain a better understanding of relationships among age, gender, educational level, financial adequacy, personal resources, internet use, and stress and prepare a more complex moderation analysis.

Second, to replicate the findings by Hobfoll and Lilly [6] and Hobfoll [21] regarding the stronger effect of resource losses than resource gains on stress, we conducted simple regression using participants’ characteristics and personal resource variables as predictors. Moreover, to explore the moderating effect of resource gains on the relationship between resource losses and stress levels, we conducted moderation analyses using PROCESS (version 3.5) by Hayes [63], model 1. These analyses allowed for the examination of how the relationship between resource losses and stress levels varied depending on the level of resource gains reported by participants (ie, effect of the 2-way interaction gains × losses on stress).

Finally, to determine whether internet use influenced the moderation effect of resource gains on the association between resource losses and stress, a moderation analysis was performed using model 3 on PROCESS (ie, version 3.5 for SPSS by Hayes [63]). This method allowed for the testing of the effect of the triple interaction, internet use × gains × losses, on stress. In all moderation analyses, we controlled for age, gender, educational level, and financial adequacy.
In each regression and moderation analysis, we divided the analysis into 2 sets: the first set included the domain-general resources as general gains and losses, and the second set included the gains and losses of domain-specific resources such as social, cognition, and self-efficacy.

Unstandardized coefficients and 95% CIs were reported. The level of statistical significance was set at $P < .05$. All analyses were performed using SPSS (version 26; IBM Corp).

**Results**

**Mean-Level Comparisons**

Mean-level tests were conducted to determine the age-related differences in stress and personal resources (Table 3).

Table 3. Younger and older adults' mean stress and personal resource variables (ie, questionnaire scores).

<table>
<thead>
<tr>
<th></th>
<th>Younger adults (n=275), mean (SD)</th>
<th>Older adults (n=224), mean (SD)</th>
<th>Test for mean-level differences (N=499)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$F$ test ($df$)</td>
</tr>
<tr>
<td><strong>Stress</strong></td>
<td>2.47 (0.70)</td>
<td>1.99 (0.69)</td>
<td>58.45 (1, 497)</td>
</tr>
<tr>
<td><strong>Resources</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Losses</strong></td>
<td>1.54 (0.63)</td>
<td>1.82 (0.76)</td>
<td>14.97 (1, 497)</td>
</tr>
<tr>
<td>Social</td>
<td>1.54 (0.63)</td>
<td>1.60 (0.72)</td>
<td>1.07 (1, 497)</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>1.74 (0.92)</td>
<td>1.90 (0.93)</td>
<td>3.62 (1, 497)</td>
</tr>
<tr>
<td>Cognition</td>
<td>1.47 (0.73)</td>
<td>1.92 (0.88)</td>
<td>39.27 (1, 497)</td>
</tr>
<tr>
<td><strong>Gains</strong></td>
<td>2.31 (0.97)</td>
<td>1.82 (0.95)</td>
<td>32.72 (1, 497)</td>
</tr>
<tr>
<td>Social</td>
<td>2.57 (1.04)</td>
<td>2.00 (1.02)</td>
<td>37.37 (1, 497)</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>2.41 (1.14)</td>
<td>1.82 (1.06)</td>
<td>34.79 (1, 497)</td>
</tr>
<tr>
<td>Cognition</td>
<td>2.03 (1.08)</td>
<td>1.67 (0.99)</td>
<td>15.57 (1, 497)</td>
</tr>
<tr>
<td>Internet use</td>
<td>4.91 (0.23)</td>
<td>3.35 (1.38)</td>
<td>342.74 (1, 497)</td>
</tr>
</tbody>
</table>

**Correlation Analysis**

We performed correlational analyses to highlight relationships among sociodemographics, independent variables (domain-general and domain-specific resources and internet use), and stress (Table 4; for a complete table of correlations, see Multimedia Appendix 2).

In the total sample group, age, gender, and financial adequacy were correlated with stress; being female, having less money than needed, and being younger were associated with higher stress. Moreover, resource losses were related to higher stress. Specifically, higher stress was associated with more social, cognition, and self-efficacy losses. Participants who used the internet more often reported higher levels of stress.

The results showed that older adults were less stressed than younger adults (1.99 vs 2.47; $P < .001$; ie, questionnaire scores). Older adults had more losses compared to younger adults (1.82 vs 1.54; $P < .001$; ie, questionnaire scores), whereas they presented less resource gains (1.82 vs 2.31; $P < .001$; ie, questionnaire scores). More specifically, the analyses conducted on domain-specific resources (ie, social, self-efficacy, and cognition) revealed that the resource losses in the social, self-efficacy, and cognitive domains tended to increase with age, whereas the social, self-efficacy, and cognitive resource gains tended to decrease with age. Finally, older adults reported less internet use than younger adults (3.35 vs 4.91; $P < .001$; ie, questionnaire scores).

In the separate age group analyses, we found that higher levels of stress were associated with higher levels of resources losses and, more specifically, social, cognitive, and self-efficacy losses in both younger and older adults. Moreover, correlations in the younger adult group revealed that being female and having less money than needed were associated with higher stress. Furthermore, higher levels of domain-general resource gains in younger adults were associated with stress, and domain-specific resource gains such as cognitive and self-efficacy were related to being less stressed. Concerning the older adult group, analyses revealed that having a higher level of education was associated with lower stress. Finally, while the use of the internet was not associated with stress scores among the younger participants, it was negatively associated with stress among the older participants, indicating that more internet use was linked to lower stress levels.
Moderation Analyses

Overview

Regarding the profiles of internet use between younger and older adults, we found notable differences. Specifically, younger adults exhibited uniformly high use rates (median 5.00, range 3.38-5.00), whereas older adults demonstrated a considerable range in their internet use behavior, from very low to very high (median 3.56, range 1.00-5.00). To evaluate the moderating effect of age on the link between resource loss, resource gain, and internet use on stress, we initially tested a model including the age group variable in a quadruple interaction term (loss × gain × internet use × age groups; not presented). However, no significant interaction was found (β = –.02; P = .95), possibly due to the complexity of the interaction term and the unequal variances in internet use across age groups. Indeed, the assumption of homogeneity of variances for internet use was not respected (Levene test: F₁,₅₀₂ = 536.63; P < .001), showing a difference in variances across groups. On the basis of these findings but also on previous research in the field suggesting that internet use may influence the level of stress, we decided to present the following analysis separately for younger and older individuals (although the results should be interpreted with caution). This methodological approach was crucial for understanding how internet use influenced the relationship between resource loss and gain and stress, allowing for an exploration of use trajectories characteristic of younger and older adults without the confounding influence of the homogeneous high internet use found in the younger cohort.

Domain-General Resources

First, simple regression (model 1; Multimedia Appendix 3) revealed that resource indicators accounted for 32% and 20% of the individual differences in stress levels in younger and older adults, respectively. Similarly, for younger and older adults, the strongest predictor was resource losses (younger adults: β = .45 and P < .001; older adults: β = .37 and P < .001) followed by resource gains (younger adults: β = –.20 and P < .001; older adults: β = –.13 and P = .01), suggesting that individuals who experienced higher resource losses and lower gains tended to feel more stressed as compared to those with lower levels of losses and higher levels of gains.

Second, we found a significant 2-way interaction (model 2; Multimedia Appendix 3 and Figure 1) between resource gains and resource losses on levels of stress as the dependent variable in both younger and older adults, confirming our second hypothesis (β = .26 and β = .16 for younger and older adults, respectively). Specifically, individuals with high levels of resource gains who also reported high levels of losses felt less stressed compared to individuals with lower levels of resource gains and higher levels of resource losses.

Finally, internet use influenced the moderation effect of resource gains on the relationship between resource losses and stress, as seen in a significant 3-way interaction among internet use, gains, and losses in older adults (model 3; Multimedia Appendix 3; β = .09). More specifically, compared to the second model, the third model presented an increase in the index of adjustment (ie, a change in explained variance) of 0.05 (ΔR² = 0.05; P < .002). The 3-way interaction was not significant in younger adults (β = .22).

To further the understanding of the 3-way interaction in older adults, we examined the conditional effects of resource losses at 1 SD above (+1 SD) and 1 SD below (–1 SD) the mean scores of resource gains (first moderator) and internet use (second moderator; Table 5 and Figure 2). As the 3-way interaction (losses × gains × internet use) was not significant in younger adults, we reported only the conditional effects of older adults. The results showed significant moderation effects of resource gains and internet use on the relationship between

| Table 4. Significant Pearson correlations (r) between participant characteristics and personal resources and stress levels for the total sample (younger and older adults; N=510). |
|-----------------|-----------------|-----------------|-----------------|
|                 | Younger adults (n=280) | Older adults (n=230) | Total sample |
|                 | r                 | P value          | r               | P value          | r               | P value          |
| Age             | –0.01             | .81              | 0.02            | .81              | –0.32           | <.001            |
| Gender          | 0.21              | <.001            | 0.11            | .08              | 0.20            | <.001            |
| Educational level | –0.03            | .59              | –0.25           | <.001            | 0.08            | .08              |
| Financial adequacy | 0.28             | <.001            | 0.12            | .08              | 0.25            | <.001            |
| Losses          | 0.42              | <.001            | 0.36            | <.001            | 0.30            | <.001            |
| Social          | 0.21              | <.001            | 0.21            | .002             | 0.18            | <.001            |
| Self-efficacy   | 0.49              | <.001            | 0.40            | <.001            | 0.40            | <.001            |
| Cognition       | 0.30              | <.001            | 0.33            | <.001            | 0.20            | <.001            |
| Gains           | –0.21             | <.001            | –0.01           | .87              | –0.03           | .49              |
| Social          | –0.08             | .18              | –0.01           | .81              | 0.04            | .41              |
| Self-efficacy   | –0.28             | <.001            | –0.01           | .85              | –0.07           | .12              |
| Cognition       | –0.18             | .003             | –0.00           | .97              | –0.04           | .35              |
| Internet use    | –0.05             | .39              | –0.17           | .01              | 0.12            | .005             |
domain-general losses and stress. Specifically, significant moderation effects were observed only in cases in which individuals reported high levels of gains and high levels of internet use ($\beta=.37$), as well as in cases in which individuals reported low levels of gains and high levels of internet use ($\beta=.55$) and low levels of both gains and internet use ($\beta=.66$). These results suggest that the use of the internet reinforces the buffering effect of gains on the relationship between resource losses and stress.

According to the third hypothesis, we conducted an additional analysis to examine the differences between individuals with high internet use and those with low internet use in relation to their gains and losses (Table 6). As the 3-way interaction (losses $\times$ gains $\times$ internet use) was not significant in younger adults, we reported only the conditional effects of older adults. The results showed that internet use effects (ie, differences between participants with a higher internet use and participants with a lower internet use) were marginally significant only in cases in which individuals reported high levels of gains and low levels of losses ($\beta=-.21$).

**Figure 1.** Mean stress depending on the level of gains (high vs low) and losses (high vs low) in younger and older adults. The slopes’ values represent the coefficients and 95% CIs.

**Table 5.** Conditional effects of domain-general resource losses (ie, independent variable) at +1 SD and –1 SD of the mean scores of gains and internet use (ie, moderators) in older adults.

<table>
<thead>
<tr>
<th></th>
<th>$\beta$ coefficient (SE; 95% CI)</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Losses at –1 SD of gains and –1 SD of internet use</td>
<td>0.66 (0.12; 0.34 to 0.89)</td>
<td>.001</td>
</tr>
<tr>
<td>Losses at –1 SD of gains and +1 SD of internet use</td>
<td>0.55 (0.14; 0.28 to 0.83)</td>
<td>.001</td>
</tr>
<tr>
<td>Losses at +1 SD of gains and –1 SD of internet use</td>
<td>0.04 (0.17; –0.30 to 0.37)</td>
<td>.84</td>
</tr>
<tr>
<td>Losses at +1 SD of gains and +1 SD of internet use</td>
<td>0.37 (0.08; 0.20 to 0.53)</td>
<td>.001</td>
</tr>
</tbody>
</table>

**Figure 2.** Mean stress depending on the level of gains (high vs low), internet use (high vs low), and losses (high vs low) in older adults.
Among younger adults, interaction effect was found to be significant for different types of resource gains on the relationship between resource losses and stress levels in both younger and older adults, as indicated by the 2-way interaction effect (internet use × losses). Specifically, younger individuals who reported high losses in social and self-efficacy resources and had higher internet use experienced more stress compared to their counterparts with lower internet use. At the same time, younger adults with low levels of social and self-efficacy losses and higher internet use exhibited lower stress levels than those with lower internet use.

In the case of older adults, a significant interaction effect was observed specifically for cognition resources ($β_{\text{Cognition}}=0.08, 95\% \text{ CI } 0.00-0.16$). Older individuals who experienced fewer losses in cognition and had higher levels of internet use reported lower levels of stress than individuals with lower internet use. In addition, a significant 2-way interaction effect between internet use and gain was marginally significant ($β_{\text{Cognition}}=0.09, 95\% \text{ CI } 0.00-0.09, 95\% \text{ CI } -0.21 \text{ to } 0.02$). Older adults who had high gains in cognition resources and high levels of internet use exhibited lower levels of stress than older adults with lower internet use.

Finally, in older adults, internet use influenced the moderating effect of self-efficacy gains on the relationship between self-efficacy losses and stress and the moderating effect of social gains on the relationship between social losses and stress. This was evidenced by the significant 3-way interactions between internet use, self-efficacy gains, and self-efficacy losses ($β_{\text{Self-efficacy}}=0.07, 95\% \text{ CI } 0.01-0.14$) but also between internet use, social gains, and social losses ($β_{\text{Social}}=0.11, 95\% \text{ CI } 0.03-0.19$). Specifically, compared to the second model, the third model presented an increase in the index of adjustment (ie, a change in explained variance) of 0.02 ($ΔR^2=0.02; P=0.03$) for the self-efficacy model and 0.03 ($ΔR^2=0.03; P=0.006$) for the social model.

It is important to note that in younger adults, regardless of the type of resources, triple interactions between internet use, self-efficacy, cognition or social gains, and self-efficacy, cognition or social losses were not significant, in line with the findings on the triple interaction of the domain-general resources (Multimedia Appendix 4). As the 3-way interaction (losses × gains × internet use) was not significant in younger adults, we reported only the conditional effects of older adults. In older adults, conditional effects analyses (Table 7 and Figure 3) revealed significant moderation effects of self-efficacy gains and internet use on the relationship between self-efficacy losses and stress. Specifically, the moderation effects were observed only when individuals reported high levels of self-efficacy gains and high levels of internet use ($β=0.29$), as well as when individuals reported low levels of self-efficacy gains and high levels of internet use ($β=0.38$) and low levels of both self-efficacy gains and internet use ($β=0.45$). Moreover, significant moderation effects were observed specifically for cognition resources ($β_{\text{Cognition}}=0.08, 95\% \text{ CI } 0.00-0.16$) and social resources ($β_{\text{Social}}=0.67, 95\% \text{ CI } 0.09-1.25$). Specifically, younger individuals who reported high losses in social and self-efficacy resources and had higher internet use experienced more stress compared to their counterparts with lower internet use. At the same time, younger adults with low levels of social and self-efficacy losses and higher internet use exhibited lower stress levels than those with lower internet use.

### Domain-Specific Resources

Similarly to the domain-general resources, we conducted 3 regression models for the domain-specific resources in both younger and older adults (self-efficacy, cognition, and social resources; Multimedia Appendix 4). The first model showed that resource losses were the strongest predictor of stress levels across different age groups regardless of the specific type of resource considered. The coefficient of losses for self-efficacy was found to be the highest in both younger and older adults, with more losses being related to higher levels of stress in both groups. Moreover, higher gains in self-efficacy, cognition, and social resources significantly predicted lower stress levels in younger adults. However, in older adults, the only significant predictor was self-efficacy gains, with more gains being associated with less stress. These findings suggest that different types of resource gains may play varying roles in shaping stress experiences across different age groups.

Moreover, the impact of resource losses on stress in younger and older adults was significantly influenced by their levels of gains, which varied depending on the type of resource. Those with high levels of self-efficacy and cognition gains tended to experience less stress when they also presented high levels of self-efficacy and cognition losses (younger adults: unstandardized coefficient $β_{\text{Cognition}}=0.21, 95\% \text{ CI } 0.08-0.34$ and $β_{\text{Self-efficacy}}=0.20, 95\% \text{ CI } 0.08-0.32$; older adults: $β_{\text{Cognition}}=0.23, 95\% \text{ CI } 0.11-0.34$ and $β_{\text{Self-efficacy}}=0.23, 95\% \text{ CI } 0.13-0.34$) compared to individuals with lower levels of resource gains who tended to experience higher levels of stress under similar circumstances (ie, younger adults: $β_{\text{Cognition}}=0.40, 95\% \text{ CI } 0.21-0.60$ and $β_{\text{Self-efficacy}}=0.40, 95\% \text{ CI } 0.30-0.50$; older adults: $β_{\text{Cognition}}=0.37, 95\% \text{ CI } 0.23-0.50$ and $β_{\text{Self-efficacy}}=0.43, 95\% \text{ CI } 0.30-0.55$). While there were no significant results regarding social resources for younger adults, older adults with high levels of social gains experienced less stress when they also presented high levels of social losses ($β_{\text{Social}}=0.18, 95\% \text{ CI } 0.04-0.32$) compared to older adults with lower levels of resource gains who experienced higher levels of stress under similar circumstances ($β_{\text{Social}}=0.45, 95\% \text{ CI } 0.23-0.67$).

This study also revealed a significant influence of internet use on the relationship between resource losses and stress levels in both younger and older adults, as indicated by the 2-way interaction effect (internet use × losses). Specifically, the interaction effect was found to be significant for different types of resources in younger and older adults. Among younger adults, a significant interaction effect was observed for self-efficacy and social resources ($β_{\text{Self-efficacy}}=0.30, 95\% \text{ CI } –0.02 \text{ to } 0.31$;

<table>
<thead>
<tr>
<th></th>
<th>β coefficient (SE; 95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internet use at –1 SD of gains and –1 SD of losses</td>
<td>–0.01 (0.06; –0.12 to 0.10)</td>
<td>.89</td>
</tr>
<tr>
<td>Internet use at –1 SD of gains and +1 SD of losses</td>
<td>–0.07 (0.08; –0.23 to 0.09)</td>
<td>.40</td>
</tr>
<tr>
<td>Internet use at +1 SD of gains and –1 SD of losses</td>
<td>–0.21 (0.12; –0.44 to 0.02)</td>
<td>.08</td>
</tr>
<tr>
<td>Internet use at +1 SD of gains and +1 SD of losses</td>
<td>–0.03 (0.06; –0.14 to 0.09)</td>
<td>.67</td>
</tr>
</tbody>
</table>

Table 6. Conditional effects of internet use (ie, independent variable) on stress at +1 SD and –1 SD of the mean scores of resource losses and gains (ie, moderators) in older adults.
effects of social gains and internet use on the relationship between social losses and stress were observed only when individuals reported high levels of social gains and high levels of internet use ($\beta=0.19$) and low levels of both social gains and internet use ($\beta=0.64$). Regarding domain-general resources, these results suggest that the use of the internet increases the buffering effect of self-efficacy and social gains on the relationship between self-efficacy and social losses and stress.

Regarding the domain-general resources, we conducted an additional analysis to examine the differences between older adults with high internet use and those with low internet use in relation to their gains and losses (Table 8). As the 3-way interaction (losses × gains × internet use) was not significant in younger adults, we reported only the conditional effects of older adults. The results showed that the internet use effects (ie, differences between participants with greater internet use and participants with lower internet use) were observed in older individuals who reported high levels of self-efficacy gains and low levels of self-efficacy losses ($\beta=-0.21$). Moreover, significant internet use effects were observed in cases in which older adults experienced low levels of social gains and high levels of social losses ($\beta=-0.18$), as well as in older individuals with low levels of social gains and high levels of social losses ($\beta=-0.17$).

Table 7. Conditional effects of domain-specific resource losses (ie, independent variable) at +1 SD and –1 SD of the mean scores of gains and internet use (ie, moderators) in older adults.

<table>
<thead>
<tr>
<th></th>
<th>$\beta$ coefficient (SE; 95% CI)</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-efficacy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Losses at –1 SD of gains and –1 SD of internet use</td>
<td>0.45 (0.07; 0.31 to 0.59)</td>
<td>.001</td>
</tr>
<tr>
<td>Losses at –1 SD of gains and +1 SD of internet use</td>
<td>0.38 (0.10; 0.18 to 0.58)</td>
<td>.001</td>
</tr>
<tr>
<td>Losses at +1 SD of gains and –1 SD of internet use</td>
<td>–0.02 (0.12; –0.26 to 0.23)</td>
<td>.88</td>
</tr>
<tr>
<td>Losses at +1 SD of gains and +1 SD of internet use</td>
<td>0.29 (0.07; 0.15 to 0.43)</td>
<td>.001</td>
</tr>
<tr>
<td><strong>Social relations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Losses at –1 SD of gains and –1 SD of internet use</td>
<td>0.64 (0.15; 0.35 to 0.93)</td>
<td>.001</td>
</tr>
<tr>
<td>Losses at –1 SD of gains and +1 SD of internet use</td>
<td>0.23 (0.16; –0.09 to 0.54)</td>
<td>.16</td>
</tr>
<tr>
<td>Losses at +1 SD of gains and –1 SD of internet use</td>
<td>–0.02 (0.14; –0.29 to 0.25)</td>
<td>.88</td>
</tr>
<tr>
<td>Losses at +1 SD of gains and +1 SD of internet use</td>
<td>0.19 (0.08; 0.02 to 0.36)</td>
<td>.02</td>
</tr>
</tbody>
</table>
Figure 3. Plots of 3-way interaction effects for the self-efficacy model (self-efficacy losses × self-efficacy gains × internet use) and the social model (social losses × social gains × internet use) in older adults.

![Graphs showing 3-way interaction effects for self-efficacy and social models](image)

Table 8. Conditional effects of internet use (ie, independent variable) at +1 SD and –1 SD of the mean scores of domain-specific resource gains and losses (ie, moderators) in older adults.

<table>
<thead>
<tr>
<th></th>
<th>β coefficient (SE; 95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Self-efficacy</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internet use at –1 SD of gains and –1 SD of losses</td>
<td>–0.01 (0.05; –0.11 to 0.09)</td>
<td>.82</td>
</tr>
<tr>
<td>Internet use at –1 SD of gains and +1 SD of losses</td>
<td>–0.06 (0.07; –0.19 to 0.07)</td>
<td>.39</td>
</tr>
<tr>
<td>Internet use at +1 SD of gains and –1 SD of losses</td>
<td>–0.25 (0.11; –0.46 to –0.04)</td>
<td>.02</td>
</tr>
<tr>
<td>Internet use at +1 SD of gains and +1 SD of losses</td>
<td>–0.05 (0.06; –0.16 to 0.06)</td>
<td>.39</td>
</tr>
<tr>
<td><strong>Social relations</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internet use at –1 SD of gains and –1 SD of losses</td>
<td>0.03 (0.06; –0.08 to 0.15)</td>
<td>.57</td>
</tr>
<tr>
<td>Internet use at –1 SD of gains and +1 SD of losses</td>
<td>–0.17 (0.09; –0.35 to 0.01)</td>
<td>.07</td>
</tr>
<tr>
<td>Internet use at +1 SD of gains and –1 SD of losses</td>
<td>–0.18 (0.08; –0.34 to –0.01)</td>
<td>.03</td>
</tr>
<tr>
<td>Internet use at +1 SD of gains and +1 SD of losses</td>
<td>–0.08 (0.06; –0.19 to 0.04)</td>
<td>.20</td>
</tr>
</tbody>
</table>

**Discussion**

**Principal Findings**

This study investigated the influence of internet use on the relationship between personal resources and stress in both younger and older adults. The findings indicated that older adults were less stressed than younger adults. Moreover, resource gains moderated the relationship between resource losses and stress, and this effect was similar in both younger and older individuals. Finally, internet use seems to act as a buffer on the dynamics between social and self-efficacy resource losses and stress, amplifying the positive influence of resource gains in reducing the adverse effects of these losses. In older
adults, internet use was beneficial as a means of dealing with losses in social and self-efficacy resources.

**Age-Associated Differences in Stress Levels**

In support of our first hypothesis, we found that older individuals reported less stress than younger adults. This finding is consistent with those of extant research that has also documented the stress-buffering effect of age among older adults. This phenomenon has been attributed to several personal factors, including cognitive and emotional processing differences between age groups [64,65], greater use of emotion regulation strategies [66,67], and greater life experience and wisdom that allow for more effective coping with stressors [68,69]. Moreover, older adults may be more skilled at regulating their emotions, which may reduce the impact of stressful events on their psychological well-being.

Specifically, stress-inducing situations are related to an increase in negative emotions, and several studies have shown that older adults tend to experience more positive and less negative emotions [70-74]. Therefore, it can be inferred that older adults may possess a greater capacity to regulate and inhibit negative emotions, leading to a reduced impact of stressful events on their psychological well-being. Older adults may further be more resilient than younger adults due to their accumulated life experience and developed coping mechanisms, including proactive problem-solving strategies, effective emotion regulation, and strong sense of personal control and self-efficacy [75-78]. For example, several studies have reported a coping shift during aging to match the constraints experienced and preserve well-being [79-82]. Older adults, who often face a range of losses associated with aging such as declining health, social network changes (eg, death of partner), and retirement, tend to exhibit a greater preference for accommodation, including emotion- and cognition-focused coping. In contrast, younger adults, who typically have fewer losses, displayed a higher preference for assimilation, including problem-focused coping, and actively sought solutions to alleviate stress, reflecting their developmental stage characterized by a stronger drive for achievement, personal growth, and the ability to confront different types of challenges. The observed shift from assimilation to accommodation [19] across the life span suggests a developmental trajectory in coping strategies, with older adults adapting their coping approaches to address the unique challenges and losses they experience. Overall, older adults may be more resilient to stress than younger adults, and their adaptive coping strategies, social support, and emotion regulation strategies may contribute to their ability to maintain psychological well-being in the face of adversity.

**Resource Gains Moderated the Relationship Between Resource Losses and Stress Levels**

We found that resource gains moderated the relationship between resource losses and stress levels, confirming our second hypothesis. In line with the COR theory [20], we found that resource gains buffered the negative impact of resource losses on stress levels. More specifically, COR theory emphasizes the significance of resource gains, which have the potential to assist individuals in restoring their resources and avoiding further depletion. The gain paradox posits that individuals who have experienced losses are more likely to recognize and appreciate resource gains. This phenomenon can be attributed to the increased awareness among individuals of the value and importance of resources as a result of experiencing losses, which in turn serves as a motivation for them to actively seek and increase those resources. Consequently, resource gains are crucial in moderating the negative relationship between resource losses and stress levels as individuals with greater resource gains possess better coping abilities and are more likely to recover from losses [21].

Replicating the COR theory by Hobfoll [20], we also found that resource gains buffered the impact of resource losses in both age groups on stress. That the same buffering effect was found in both age groups could be attributed to some universality of the losses-gains dynamic across life phases. The COR theory posits that individuals across the life span share a fundamental drive to accumulate and protect resources as a means of maintaining well-being and minimizing stress [20]. Therefore, the importance of resource gains in mitigating the negative impact of resource losses on stress levels may hold regardless of age.

**Effect of the Relationship Between Internet Use and Domain-Specific Resources on Stress in Younger and Older Adults**

In addition, this study revealed distinct patterns of interaction between effect of internet use as an external technical resource and domain-specific personal resources on stress levels among younger and older adults. Our findings suggest that the relationship between social or self-efficacy losses and stress levels in younger individuals is significantly influenced by their use of the internet. Specifically, younger individuals who reported fewer social or self-efficacy losses experienced lower levels of stress when they used the internet, indicating a buffering effect. On the other hand, those who reported higher social or self-efficacy losses exhibited increased levels of stress when they used the internet, suggesting an exacerbating effect. These results highlight the complex interplay among social or self-efficacy losses, internet use, and stress levels in younger individuals. It appears that the internet may serve as a supportive resource for individuals with fewer social losses, providing them with a means for social connection and support [83]. Furthermore, the internet may serve as a resource for providing access to information, support, and opportunities for skill development, which can bolster self-efficacy beliefs and resilience and promote adaptive coping strategies [84-88].

However, for those experiencing higher social or self-efficacy losses, the internet may exacerbate stress. Indeed, the losses in social resources, such as social contacts, are associated with higher feelings of loneliness [89], which contribute to the development of excessive internet use, commonly referred to as internet addiction [90,91]. This pattern of excessive internet use, driven by the absence of social support and challenges in communication and in emotion identification and regulation, is linked to higher levels of stress [92,93]. This suggests that individuals experiencing significant social losses may increase their use of the internet as a compensatory mechanism to mitigate the impact of these losses, resulting in increased stress...
levels. Moreover, previous research has documented a generational-situated use of the internet, with younger adults using it for leisure activities whereas older individuals’ preferred use of the internet is to facilitate the realization of daily activities such as medical consultations [94,95].

In addition, this study highlights an interesting pattern regarding the relationship among internet use, cognitive losses, and stress levels in older adults. Specifically, older adults with lower levels of cognitive losses who engaged in internet use experienced lower levels of stress compared to those who did not use the internet, suggesting that internet use may serve as a protective means against stress for older adults with fewer cognitive losses. One possible explanation can be that internet use provides opportunities to be engaged in web-based activities for cognitive stimulation, accessing information, or social interaction [8,96], which may help mitigate the negative effects of cognitive losses on stress levels. Considering that higher frequency of digital device use has been associated with fewer subjective cognitive concerns [97-99] and that cognitive losses have been associated with higher levels of stress in older adults [23,100,101], we found that older adults with less cognitive losses who used the internet more were less stressed compared to older adults who did not use the internet, suggesting a buffer effect of internet use on the relationship between cognitive losses and stress.

However, older adults with high levels of cognitive losses experienced similar levels of stress regardless of their internet use. This suggests that the influence of high cognitive losses on stress may remove any potential benefits derived from internet use. It is possible that older adults with high cognitive losses may have difficulties using the internet effectively due to their subjective cognitive losses, which could be explained by the digital distraction hypothesis [102-104]. According to this hypothesis, increased engagement with technology may have detrimental consequences for cognitive processes, manifesting as executive dysfunction characterized by heightened distractibility, superficial cognitive processing, and difficulties in task organization and completion. In addition, technology reliance may contribute to increased forgetfulness by undermining the natural memory systems used for tasks such as navigation or recalling personal information such as phone numbers [102,104]. The detrimental effects of excessive digital engagement on cognitive functioning may override any potential benefits of internet use for stress reduction in older adults with higher cognitive losses.

**Internet Use Moderated the Relationship Between Resource Losses and Gains and Stress Levels in Older Adults**

The final hypothesis of our study, which examined the influence of internet use on the relationship between resources and stress, was confirmed for older adults. More specifically, older adults with fewer losses in self-efficacy and social resources and greater gains in these domains experienced lower levels of stress when they engaged in more internet use. This suggests that the internet can be considered as an “amplifier” of the positive effects of resource gains, particularly in terms of self-efficacy and social resources.

Previous studies have reported that internet use by older adults has been associated with decreased loneliness and depression; better social connectedness, self-esteem, and cognitive functioning [105,106]; and improved self-efficacy, self-control, self-determination, and skill development [107-111]. For example, the study by Karavidas et al [110] examined the association between internet use, self-efficacy resources, and life satisfaction among older adults. The results revealed a positive correlation between internet use and life satisfaction. This relationship was mediated by self-efficacy resources, indicating that increased internet use among older individuals was associated with the development of higher self-efficacy skills, which in turn contributed to an improved overall quality of life. The findings suggest that frequent internet use may serve as a platform for older adults to develop and enhance their self-efficacy, leading to greater life satisfaction. Similarly, Chaumon et al [107] found that older adults with functional loss living in long-term care institutions showed a positive impact of internet use on self-sufficiency, self-efficiency, and psychological empowerment [112].

Moreover, the internet facilitates stronger social connections and easier access to social networks, such as through engaging in web-based conversations with new contacts or actively participating in web-based social events [76,105,107,113]. For example, White et al [105] presented a randomized controlled trial to investigate the psychosocial impact of providing internet training and access to older adults. The study involved a sample of older individuals who were randomly assigned to either an intervention group, which received internet training, or a control group that did not receive any intervention. The results showed significant improvements in several psychosocial factors among the intervention group compared to the control group. Specifically, older adults who received internet training reported increased social support, higher levels of social engagement, reduced feelings of loneliness, and enhanced subjective well-being. These findings suggest that providing older adults with internet training and access can have positive effects on their psychosocial well-being.

In line with previous findings, we found that internet use can support the buffer effects from gains in self-efficacy and social resources when individuals have low levels of losses in each type of resource. Moreover, internet use can also substitute the effects of gains in older adults with low gains and high losses. More specifically, our findings demonstrated that individuals who reported high losses and low gains in social resources experienced lower levels of stress when they engaged in more internet use. This can be attributed to the compensatory role of the internet in filling the gaps caused by the limited gains in social resources. Several studies have provided evidence supporting the notion that the internet can compensate for losses in social resources among individuals [114-118]. Older adults who experience a decline in face-to-face social interactions due to factors such as retirement or physical limitations can benefit from web-based social networking platforms. For example, Khoo and Yang [116] conducted a study that examined the impact of social media use on the perception of social support among middle-aged and older adults. The researchers found that using social media platforms for interactions with broader...
social networks such as friends was as beneficial as using them to connect with family members in terms of enhancing social support.

**Limitations**

This study entails certain limitations that warrant consideration. These pertain both to the representativeness of the recruited sample and to the measures used therein [119]. On the one hand, the results should be interpreted cautiously due to potential sampling error. Indeed, stemming from a nonprobabilistic sampling approach, the characteristics of the participants who voluntarily engaged in the study may also influence the variables of interest [119,120]. For instance, participants’ income could influence their willingness to participate in surveys [121], their resource losses and gains, and their internet use [122]. Consequently, the sampling strategy used (ie, snowball volunteer sampling) likely did not capture older individuals marked by resource losses and sociodemographic characteristics (eg, isolation and low socioeconomic status) that are particularly stressful.

On the other hand, the nature of the measures used may introduce limitations. The results should be interpreted with caution when examining age-related differences in the role of internet use. Reflecting on the issue of causality, it is important to consider how the cross-sectional nature of our study limits our ability to ascertain changes in internet use and its effects over time, especially across different age groups [52]. While we identified associations among changes in personal resources, stress levels, and internet use, these findings are not sufficient to establish a causal link or to delineate the temporal evolution of internet use’s impact on stress and personal resources. This caution extends to interpreting the dynamics of internet use across the life span, where cross-sequential research is essential for distinguishing between the effects of aging and those attributable to cohort-specific experiences or generational differences [94].

Moreover, the decision not to include age group as a factor in an interaction analysis was based on significant differences in internet use between younger and older adults. Preliminary findings indicated uniform high use among younger participants, which could confound nuanced age-related interactions with internet use and stress outcomes. Therefore, analyses were conducted separately for each age group to accurately capture distinct use patterns, especially among the older adults who demonstrated a broader range of internet behaviors. This approach helped avoid the confounding effects of uniform use in younger adults. However, it limited the exploration of broader age-related dynamics, potentially affecting a comprehensive understanding of how age influences the relationship between internet use and stress.

In addition, it is essential to recognize that this study’s focus primarily lay on the frequency of internet use, measured equivalently for both younger and older adults. However, this approach disregards the plausible generational disparities in internet use patterns, as noted in previous research [94,123], which might inadvertently introduce errors in measurement [119]. Notably, the lack of statistical significance concerning the impact of internet use on stress in younger adults, a group extensively engaged with the internet, might potentially be attributed to the distinct “youthful” internet use styles, as proposed by Boullier [94]. Indeed, the presence of measurement errors, such as those arising from questionnaire elaboration, can introduce challenges such as ceiling and floor effects [119,124], impeding the identification of statistically significant differences between groups [125]. Moreover, future studies should develop measures to better capture interindividual differences in internet use in younger individuals to further investigate the multifaceted relationship between internet use and different age groups.

Moreover, the consideration of a singular internet use style in measurement limits the possibility of identifying the styles that may be most beneficial for each age group during resource losses. Indeed, the beneficial role of internet use derives from the meanings attributed to its use and the opportunities it offers to address needs [126-128]. Consequently, it would be of great interest for future studies to focus on identifying favorable internet use styles for stress reduction in a context of personal resource loss across different age groups.

**Conclusions**

This study examined the effect of internet use on the relationship between personal resources and stress. The findings revealed a nuanced understanding of how digital engagement can serve as a buffer against stress, particularly among older adults who experienced resource losses and gains. Specifically, older adults who frequently used the internet reported less stress when exposed to high levels of both resource losses and gains compared to their counterparts with lower levels of internet engagement. This underscores the importance of internet use in mitigating stress among older adults, highlighting the potential of digital tools in promoting well-being in older populations.

This study’s novel contribution lies in its empirical support for the beneficial role of internet use among older adults within the framework of the COR theory by Hobfoll [20]. By demonstrating that internet use can moderate the effects of resource losses and gains on stress, this research provides valuable insights for developing targeted interventions aimed at leveraging technology to support well-being. Understanding these dynamics will help researchers, practitioners, and policy makers recognize the role of the internet as a facilitator of resource gains and a compensatory mechanism for social deficits in older adults. Encouraging and supporting older adults in using the internet can promote access to valuable resources and enhance their self-efficacy and social connections, ultimately contributing to improved stress management and, more generally, well-being.
Acknowledgments

This work was supported by the Swiss National Centre of Competence in Research LIVES –Overcoming Vulnerability: A Life-Course Perspective, which is financed by the Swiss National Science Foundation (grant 51NF40-185901). The authors are grateful to the Swiss National Science Foundation for its financial assistance.

Authors' Contributions

AR was responsible for the study design; data collection, analysis, and interpretation; and writing of the original draft. PM contributed to data analysis and interpretation and the writing, review, and editing of the draft. CL contributed to study design and review. DSJ contributed to review and editing of the draft. All authors agreed to the submission of the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Checklist for Reporting Results of Internet E-Surveys.
[DOCX File, 21 KB - aging_v71e52555_app1.docx ]

Multimedia Appendix 2
Pearson correlations (r) between the main variables.
[DOCX File, 28 KB - aging_v71e52555_app2.docx ]

Multimedia Appendix 3
Unstandardized regression coefficients for domain-general resources in younger and older adults.
[DOCX File, 20 KB - aging_v71e52555_app3.docx ]

Multimedia Appendix 4
Unstandardized regression coefficients for domain-specific resources (self-efficacy, cognition, and social relations) in younger and older adults.
[DOCX File, 27 KB - aging_v71e52555_app4.docx ]

References

1. Two-thirds of the world's population uses the Internet, but 2.7 billion people remain offline. International Telecommunication Union. URL: https://www.itu.int/itu-d/reports/statistics/2022/11/24/ffc2-internet-use/ [accessed 2022-07-09]


67. Reed AE, Carstensen LL. The theory behind the age-related positivity effect. Front Psychol 2012;3:339 [FREE Full text] [Medline: 23060825]


120. Groves RM. Nonresponse rates and nonresponse bias in household surveys. Public Opin Q 2006;70(5):646-675 [FREE Full text] [doi: 10.1016/j.piq.2006.03.001]


Abbreviations

COR: Conservation of Resources
Abstract

Background: The occurrence of the COVID-19 pandemic demanded fast changes in the delivery of health care. As a result, significant growth in the use of telemedicine services occurred. Research, especially from nationally representative German samples, is needed to better understand determinants of telemedicine use.

Objective: The purpose of this study was to identify determinants of telemedicine service use among middle-aged and older adults during the COVID-19 pandemic in Germany.

Methods: Cross-sectional, nationally representative data were taken from the German sample of the Survey of Health, Ageing and Retirement in Europe (SHARE). The German Corona Survey 2 (n=2039), which was conducted between June and August 2021, was used for this study. Reporting experience with remote medical consultations during the COVID-19 pandemic served as the outcome measure. Associations with socioeconomic, psychological, social, health-related, and COVID-19–related determinants were examined using multiple Firth logistic regressions.

Results: Psychological factors including feeling nervous, anxious, or on edge (odds ratio [OR] 1.61, 95% CI 1.04-2.50; \( P = .03 \)), feeling sad or depressed (OR 1.62, 95% CI 1.05-2.51; \( P = .03 \)) and feelings of loneliness (OR 1.66, 95% CI 1.07-2.58; \( P = .02 \)) were positively associated with telemedicine use. Moreover, forgoing medical treatment because of being afraid of being infected by SARS-CoV-2 (OR 1.81, 95% CI 1.10-2.97; \( P = .02 \)) and describing limitations because of a health problem as severe were positively associated with the outcome (OR 2.11, 95% CI 1.12-4.00; \( P = .02 \)). Socioeconomic and social factors were not significantly associated with telemedicine use in our sample.

Conclusions: Middle-aged and older individuals in Germany seem to use telemedicine services according to psychological needs and health limitations. Especially when psychological symptoms are experienced, telemedicine seems to be a promising service option in this age group. Future research is needed to confirm these initial findings in postpandemic circumstances.

JMIR Aging 2024;7:e50938 doi:10.2196/50938

KEYWORDS
telemedicine; telehealth; digital health; service use; COVID-19

Introduction

Telemedicine has been a big part of the digital transformation of the health care sector. Multiple definitions of telemedicine have been introduced in the past [1]. The World Health Organization (WHO) Global Observatory for eHealth [2] identified the four key characteristics of telemedicine: (1) its purpose is to provide clinical support, (2) it is intended to overcome geographical barriers, connecting users who are not in the same physical location, (3) it involves the use of various types of information and communication technology, and (4) its goal is to improve health outcomes. Therefore, telemedicine includes synchronous (eg, videoconferencing, telephone) as well as asynchronous (eg, mobile apps) health services, which are delivered via electronic devices.

Telemedicine is a presumably promising method to provide health care in the future, as it can improve access to care, save costs, and close treatment gaps [3,4]. For example, it could be a potentially valuable tool when dealing with future shortages of physicians as well as the increased demand for health care services caused by population aging; underserved rural areas can also easily be reached through telemedicine [3,5]. In past research, telemedicine was found to be an effective and cost-effective service delivery model that can be equal to in-patient visits in a variety of specialties [3,6-9]. Additionally, practitioners (eg, physicians and psychotherapists), as well as patients with conditions covered by various specialties, were found to be greatly satisfied with this form of service [8-10]. Despite the clear advantages of telemedicine, it has not yet been widely implemented. Implementations have often been decelerated by limitations regarding reimbursement, as well as
clinical, legal, cost, and social issues [11,12]. Telemedicine rates of use were gradually increasing over the previous years but remained at a low level [13].

The occurrence of the COVID-19 pandemic demanded fast changes in the delivery of health care. COVID-19 caused a major public health burden globally and it was essential to reduce in-person contacts to decrease further spreading of the virus [14]. Therefore, many nonessential appointments with physicians were canceled or postponed [15]. In a German population-based sample, this was the case for about one-third of respondents after containment measures were implemented in March 2020 [16]. Moreover, overall use of health care decreased in the first months of the pandemic, a phenomenon that was also observed in Germany [17].

Telemedicine appeared to be a key solution to major pandemic challenges. To facilitate the transformation to telemedicine, changes in infrastructure, reimbursement, and legal conditions were made worldwide. In Germany, legal efforts for the digital transformation of the health care system had already been made in November 2019 with the Digital Healthcare Act (Digitale-Versorgung-Gesetz) [18]. Consequently, it was easier for physicians to prescribe, deliver, and bill for telemedicine, and also for patients to use the services. In response to the pandemic, additional teleconsultation services were developed and regulations concerning video consultations were adapted [19,20]. Thus, telemedicine rates of use increased tremendously [21]. For example, the trend report of the Central Research Institute of Ambulatory Care in Germany reported that the number of video consultations increased from 743 conducted in December 2019 to 302,180 conducted in December 2020 [21]. Especially when there was limited access to in-person medical appointments during the pandemic in Germany, telemedicine services seemed to be a frequently used alternative, and satisfaction with the services was found to be moderately high [22,23]. Even though telephone services were used most frequently in Germany during the pandemic, a sharp increase was observed in the use of video consultations [24,25].

Besides technological, financial, organizational, and legal aspects, patient acceptance seems to be a crucial factor for the successful implementation of telemedicine services [26]. Patient characteristics that have been found to be associated with telemedicine use in past systematic reviews include, for example, age, gender, education, marital status, health status, and prior experience with computers and health technology [27,28]. A preliminary selective review of large-scale studies that were conducted during the pandemic in the United States found that telemedicine rates of use were higher among patients from urban areas, areas with greater broadband availability, and areas with higher prepandemic levels of telehealth use [29]. Moreover, being White; speaking English as first language; having health insurance, higher income, and greater disease burden; and being middle-aged were associated with greater use [29]. Nevertheless, more studies examining the use of telemedicine services and associated patient characteristics during pandemic times are needed.

Studies with samples from Germany, where telemedicine played a major role and was frequently used during the pandemic, are especially scarce. Few studies have looked at different German samples during the pandemic. While some of these studies examined large, nationally representative or quota-based samples [22,23,30], other studies only included convenience or smaller selective samples [31-33]. These studies identified potential socioeconomic (male sex, younger age, higher or lower education, living with a partner in the same household, having children younger than 18 years), psychosocial (increased loneliness, increased life satisfaction, severe psychological distress, frequent social isolation, lack of company), health-related (poor self-rated health), experience-related (higher electronic literacy, past use of telemedicine) and COVID-19–related (higher perceived severity of COVID-19 infection, having had COVID-19 infection, subjective COVID-19–related challenges, COVID-19–related cognitive preoccupation, anxiety, and worries) determinants that were positively associated with actual telemedicine use during the pandemic [22,23,30-33]. However, more studies including large nationally representative samples from Germany are needed to secure these initial findings. Moreover, the different categories of determinants, which were only partly included in single studies (eg, psychosocial or COVID-19–related determinants), should be explored further.

Middle-aged and older individuals represent the largest age group in the German population [34]. Considering population aging, the proportion of middle-aged and older adults will grow even further in the near future in Germany. Due to their increased need for health care (due to, eg, chronic conditions, frailty, and cognitive or functional decline) and potentially limited mobility, these age groups represent a major target group for future telemedicine services. Nevertheless, they seem to use telemedicine less often than other age groups [28]. Although past systematic reviews found that telemedicine is an effective and feasible service delivery model in older adults, it was also stated that further research was required to determine how services could be adapted to the individual needs of older patients [35-37]. Better understanding the telemedicine use behavior of middle-aged and older individuals could significantly contribute to increased use, as well as widespread acceptance and satisfaction with future telemedicine services. Therefore, this study aimed to explore determinants of patient use of telemedicine services in a nationally representative sample of middle-aged and older individuals during the COVID-19 pandemic in Germany.

Methods

Sample

Cross-sectional data for this study were taken from the Corona Survey 2 [38] of the Survey of Health, Ageing and Retirement in Europe (SHARE) study [39]. SHARE is a multidisciplinary and cross-national panel study that explores health, social, economic, and environmental policies in individuals aged 50 years or older and their partners (regardless of age) from 26 European countries, Switzerland, and Israel. Starting in 2004, SHARE has so far conducted 8 waves. In each wave, new respondents are added to the sample to compensate for attrition. In response to the global COVID-19 pandemic, a special Corona
Survey, to examine the health-related and socioeconomic impact of COVID-19, was introduced in June 2020. In the course of this survey, the usual computer-assisted personal interviews were replaced by telephone-administered interviews.

Participation rates for waves 1 to 8 and the Corona Survey 1 have been provided by SHARE [40,41]. According to Bergmann et al [40], these rates increased over time and reflect high overall panel stability. The final rates for the Corona Survey 2 are not available yet; however, an average retention rate of 86% (excluding recovery of respondents) was confirmed by SHARE user support. Due to the fact that SHARE assessed telemedicine service use for the first time in the Corona Survey 2, only data from this survey, which was conducted from June until August 2021, were included in our analyses. Moreover, only the German subsample (n=2039) was considered, which was done to promote comparability among participants due to existing heterogeneity regarding characteristics of health care systems, telemedicine regulations, and telemedicine use across the different countries [20,42,43]. For example, whereas countries such as Denmark, Italy, and Germany are described as advanced in telemedicine use trends, countries like Poland, Portugal, or Slovakia are still developing in the telemedicine field since the pandemic [42].

Ethical Considerations

Verbal informed consent was collected from all individuals that participated in the telephone-administered interviews for the Corona Survey 2. The SHARE project has been repeatedly reviewed and approved by the ethics committee of the University of Mannheim (waves 1-3) and the Ethics Council of the Max Planck Society (waves 4-9; most recently in June 2021 with the ethics approval number 2021_24).

Dependent Variable

In the SHARE Corona Survey 2, telemedicine service use by middle-aged and older adults during the pandemic was assessed with a metric variable: “Since the outbreak of Corona, how many remote medical consultations over the phone, computer, or any other electronic means, did you have, if any, with or without video?” Therefore, this study did not focus on one specific form of telemedicine or patient group and included consultations on online platforms (eg, video calls), as well as telephone appointments. The response format in the original questionnaire was numerical (ie, the number of experienced telemedicine consultations). For the sake of this analysis, this item was dichotomized (1=one or more remote medical consultations since the outbreak of the COVID-19 pandemic; 0=no use) because of small case numbers.

Independent Variables

Independent variables were chosen in line with former research and based on theoretical considerations. Previous systematic reviews identified mostly socioeconomic (eg, sex, age, education, relationship status, area lived in) and health (eg, disease burden, psychological symptoms) determinants of telemedicine use [27-29]. Moreover, we considered the pandemic context, including the pandemic and social consequences, when choosing independent variables. Therefore, socioeconomic, psychological, social, health-related, and COVID-19–related factors were taken into account to explore their potential relationships with telemedicine service use. Socioeconomic factors that were included were sex, age, area lived in (big city, suburbs or outskirts of a big city, large town, small town, rural area or village), living with a partner in the same household (yes or no), employment status (retired, employed or self-employed, unemployed, or other) and the household's ability to make ends meet regarding their total monthly income (with great difficulty or some difficulty, fairly easily, easily). Included psychological factors were feeling nervous, anxious, or on edge in the last month (yes or no), feeling sad or depressed in the last month (yes or no), and feeling lonely in the last month (yes or no). Furthermore, social factors included social and electronic contact frequency with people other than relatives (never, less than once a week, about once a week, several times a week, daily). Concerning health-related factors, having trouble sleeping recently (yes or no), the number of physical illnesses (including hip fracture; diabetes or high blood sugar; high blood pressure or hypertension; a heart attack, including myocardial infarction or coronary thrombosis; any other heart problem, including congestive heart failure; chronic lung diseases, such as chronic bronchitis or emphysema; and cancer or malignant tumor, including leukemia or lymphoma, but excluding minor skin cancers), limitations because of a health problem in usual activities (not limited; limited, but not severely; severely limited) and self-rated health (excellent, very good, good, fair, poor) were inspected. COVID-19–related factors that were included in the analyses were having received the COVID-19 vaccination (yes or no), oneself or anyone close having tested positive for COVID-19 (yes or no), forgoing medical treatment because of being afraid to become infected by SARS-CoV-2 (yes or no), and taking drugs or medicine as prevention against COVID-19 (yes or no).

Statistical Analysis

First, sample characteristics were computed. Second, Firth logistic regressions were conducted to identify determinants of telemedicine service use during the pandemic. The Firth method was used to reduce small-sample bias, considering the small case numbers for some of the variables [44]. Due to high correlations, the variables regarding psychological symptoms (including feeling nervous, anxious, or on edge; feeling sad or depressed; and feeling lonely in the last month) were entered separately into the model. For sensitivity analyses, conventional multiple logistic regressions were also performed. Moreover, we computed additional analyses with age as a categorical variable (40-64 years, 65-74 years, ≥75 years) to test for a nonlinear relationship with the outcome. Odds ratios (ORs) are presented with the 95% CI. P values were considered statistically significant at an α level of <.05. Since the number of missing values for the independent variables was very small (below 1%), we did not use imputation techniques. Small levels of missing values are usually less likely to significantly bias results [45,46]. Therefore, listwise deletion was applied. Stata (version 16.1; StataCorp) was used for all statistical analyses.
Results

Sample Characteristics
The total sample consisted of 2039 individuals. The sample characteristics for all included variables are presented in Table 1. Overall, 54.2% (1105/2039) of the sample were women. The mean age of the participants was 70.6 (SD 8.7) years. Considering telemedicine service use during the pandemic, 5.7% (115/2031) of the sample reported that they had had remote medical consultations at least once.
### Table 1. Sample characteristics (N=2039).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Telemedicine service use, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>1916 (94.3)</td>
</tr>
<tr>
<td>At least once</td>
<td>115 (5.7)</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>934 (45.8)</td>
</tr>
<tr>
<td>Female</td>
<td>1105 (54.2)</td>
</tr>
<tr>
<td><strong>Age (years), mean (SD)</strong></td>
<td></td>
</tr>
<tr>
<td>70.6 (8.7)</td>
<td></td>
</tr>
<tr>
<td><strong>Age (years), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>40-64</td>
<td>573 (28.1)</td>
</tr>
<tr>
<td>65-74</td>
<td>1083 (53.1)</td>
</tr>
<tr>
<td>≥75</td>
<td>383 (18.7)</td>
</tr>
<tr>
<td><strong>Area lived in, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Big city</td>
<td>289 (14.3)</td>
</tr>
<tr>
<td>Suburbs or outskirts of a big city</td>
<td>177 (8.8)</td>
</tr>
<tr>
<td>Large town</td>
<td>198 (9.8)</td>
</tr>
<tr>
<td>Small town</td>
<td>549 (27.2)</td>
</tr>
<tr>
<td>Rural area or village</td>
<td>807 (40)</td>
</tr>
<tr>
<td><strong>Living with partner in the same household, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1521 (74.6)</td>
</tr>
<tr>
<td>No</td>
<td>518 (25.4)</td>
</tr>
<tr>
<td><strong>Employment situation, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Retired</td>
<td>1445 (70.9)</td>
</tr>
<tr>
<td>Employed or self-employed</td>
<td>457 (22.4)</td>
</tr>
<tr>
<td>Other</td>
<td>135 (6.6)</td>
</tr>
<tr>
<td><strong>Households’ ability to make ends meet, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>With great or some difficulty</td>
<td>151 (7.4)</td>
</tr>
<tr>
<td>Fairly easily</td>
<td>599 (29.6)</td>
</tr>
<tr>
<td>Easily</td>
<td>1277 (63)</td>
</tr>
<tr>
<td><strong>Nervous, anxious, or on edge in the last month, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>524 (25.8)</td>
</tr>
<tr>
<td>No</td>
<td>1507 (74.2)</td>
</tr>
<tr>
<td><strong>Sad or depressed in the last month, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>609 (30)</td>
</tr>
<tr>
<td>No</td>
<td>1419 (70)</td>
</tr>
<tr>
<td><strong>Feelings of loneliness in the last month, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>442 (21.8)</td>
</tr>
<tr>
<td>No</td>
<td>1583 (78.2)</td>
</tr>
<tr>
<td><strong>Frequency of social contacts with nonrelatives, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>229 (11.3)</td>
</tr>
<tr>
<td>Less than once a week</td>
<td>583 (28.7)</td>
</tr>
<tr>
<td>About once a week</td>
<td>452 (22.3)</td>
</tr>
<tr>
<td>Several times a week</td>
<td>452 (22.3)</td>
</tr>
<tr>
<td>Characteristics</td>
<td>Values</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------</td>
</tr>
<tr>
<td>Daily frequency of electronic contacts with nonrelatives, n (%)</td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>253 (12.5)</td>
</tr>
<tr>
<td>Less than once a week</td>
<td>617 (30.4)</td>
</tr>
<tr>
<td>About once a week</td>
<td>505 (24.9)</td>
</tr>
<tr>
<td>Several times a week</td>
<td>504 (24.8)</td>
</tr>
<tr>
<td>Daily</td>
<td>150 (7.4)</td>
</tr>
<tr>
<td>Having trouble sleeping recently, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>686 (33.8)</td>
</tr>
<tr>
<td>No</td>
<td>1344 (66.2)</td>
</tr>
<tr>
<td>Number of physical illnesses (range 0-6), mean (SD)</td>
<td>1.1 (1)</td>
</tr>
<tr>
<td>Health limitations, n (%)</td>
<td></td>
</tr>
<tr>
<td>Severely limited</td>
<td>310 (15.2)</td>
</tr>
<tr>
<td>Limited, but not severely</td>
<td>744 (36.6)</td>
</tr>
<tr>
<td>Not limited</td>
<td>981 (48.2)</td>
</tr>
<tr>
<td>Self-rated health, n (%)</td>
<td></td>
</tr>
<tr>
<td>Excellent</td>
<td>82 (4)</td>
</tr>
<tr>
<td>Very good</td>
<td>320 (15.7)</td>
</tr>
<tr>
<td>Good</td>
<td>839 (41.2)</td>
</tr>
<tr>
<td>Fair</td>
<td>629 (30.9)</td>
</tr>
<tr>
<td>Poor</td>
<td>166 (8.2)</td>
</tr>
<tr>
<td>Received COVID-19 vaccine, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>189 (9.3)</td>
</tr>
<tr>
<td>No</td>
<td>1847 (90.7)</td>
</tr>
<tr>
<td>Self or anyone close tested positive for COVID-19, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>592 (29.1)</td>
</tr>
<tr>
<td>No</td>
<td>1441 (70.9)</td>
</tr>
<tr>
<td>Forgone medical treatment since COVID-19 pandemic, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>225 (11.1)</td>
</tr>
<tr>
<td>No</td>
<td>1810 (88.9)</td>
</tr>
<tr>
<td>Took drugs/medicine as prevention against COVID-19, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>45 (2.2)</td>
</tr>
<tr>
<td>No</td>
<td>1990 (97.8)</td>
</tr>
</tbody>
</table>

**Regression Analysis**

The analytic sample for the Firth logistic regressions included 1976 individuals. Results for the model, including anxiety symptoms, are presented in Figure 1 (see Multimedia Appendix 1 for detailed results for all models). Most of the included independent variables were not significantly associated with the outcome (P > .05). However, some associations were observed. Psychological factors, including feeling nervous, anxious, or on edge (OR 1.61, 95% CI 1.04-2.50; P = .03), feeling sad or depressed (OR 1.62, 95% CI 1.05-2.51; P = .03) and feeling lonely (OR 1.66, 95% CI 1.07-2.58; P = .02) in the last month were positively associated with the likelihood of telemedicine service use during the pandemic. Moreover, forgoing medical treatment because of being afraid to become infected by SARS-CoV-2 was positively associated with the outcome (OR 1.81, 95% CI 1.10-2.97; P = .02). Describing limitations because of health problems in the last 6 months as severe was also positively associated with the likelihood of telemedicine use during the pandemic (OR 2.11, 95% CI 1.12-4.00; P = .02).
In a sensitivity analysis in which Firth logistic regressions were replaced by conventional logistic regressions, similar associations were observed (see Tables S1-S3 in Multimedia Appendix 1 for detailed results). When age was included as a categorical variable in the models (see Table S4 in Multimedia Appendix 1 for detailed results), the youngest age group (40-64 years) was significantly more likely to use telemedicine services compared to the older age groups in our sample (65-74 and ≥75 years). Moreover, in these models, being employed or self-employed versus retired achieved statistical significance and was negatively associated with telemedicine use.

**Discussion**

**Key Findings**

This study aimed to identify determinants of telemedicine service use in a nationally representative sample of middle-aged and older adults during the COVID-19 pandemic in Germany. Based on data from the German sample of the SHARE Corona Survey 2, some associations of patient characteristics with telemedicine use were identified. This partly included health, psychological, and COVID-19–related factors. Socioeconomic and social determinants were not significantly associated with the outcome in this sample. So far, there has been limited...
research on determinants of telemedicine use in German samples. Our study findings thus extend our current knowledge regarding socioeconomic, social, health, psychological, and COVID-19–related determinants.

Relation to Previous Research

Whereas none of the included socioeconomic determinants were associated with telemedicine service use in our sample, some recent studies identified significant relationships. Findings regarding associations of sex and age with telemedicine use were mixed in recent studies conducted during the pandemic in Germany, with some studies observing higher rates of use in male and younger individuals [22,23,30-32]. When observing age groups in our sample, older age groups (65-74 and ≥75 years) were less likely to use telemedicine. While our study did not indicate a significant association, Hajek and König [30] observed that middle-aged and older individuals who reported living with a partner in the same household were more likely to have participated in online consultations with physicians or therapists during the pandemic. These mixed findings could potentially be explained by variations in outcomes (eg, web-based consultations vs mobile app use), samples (eg, all age groups vs only those middle-aged and older), and time frames (2020 vs 2021 vs 2022) of the different studies. This clearly highlights the need for further studies on sociodemographic determinants in German samples. Similar to our results, employment status, financial situation, and area lived in were not significantly associated with telemedicine use in other German samples during the pandemic [23,30]. However, this is in contrast to research from the United States regarding telemedicine use during the pandemic [29,47]. This contrast may be explained by a larger variation in state-specific telehealth policies before and during the pandemic [48], as well as access factors, such as possession of digital devices or availability of high speed internet [49] in the United States compared to the German samples. In contrast to Germany, health care insurance is not obligatory in the United States and additional costs arise for uninsured individuals [50], which could have contributed to telemedicine use disparities caused by socioeconomic factors in the United States [49,51-53]. Further attention should be given to the impact of socioeconomic factors on telemedicine use in future research, especially with respect to postpandemic changes and the increasing availability of in-person visits.

Our study is one of very few that has examined the association of social determinants (ie, electronic and social contact frequency) and telemedicine service use. These determinants were not significantly associated with telemedicine use in our sample. This could mean that middle-aged and older adults used telemedicine services during the pandemic based more on health factors than on reduced social contact. Nevertheless, Rauschenberg et al [33] observed that telemedicine use was higher among young individuals who reported higher perceived social isolation and lack of company during the pandemic in Germany. These contrasting findings may imply that younger individuals have used telemedicine more frequently to deal with reduced social contact during the pandemic.

Furthermore, we found that perceiving one’s limitations because of a health problem as severe was associated with telemedicine service use. This suggests that individuals with severe health limitations preferentially used telemedicine services during the pandemic. Likewise, Hajek and König [30] found a significant association of poor self-rated health and telemedicine use during the pandemic in Germany. Additionally, a positive association of disease burden and telemedicine use was observed by Harju and Neufeld [29] in large-scale US samples during the pandemic. Potential reasons for that could include the (urgent) need for treatment, limited mobility, or precautions due to high risk of severe illness from COVID-19. Patients might have used telemedicine because of health needs and lack of in-person services during the pandemic. In contrast, we found that the number of physical illnesses and self-rated health were not associated with telemedicine use during the pandemic in our sample. A potential reason for that could be that these determinants may not reflect the actual need for medicine or telemedicine services. For example, having ever received a diagnosis of physical illnesses such as hip fracture, high blood sugar, or high blood pressure does not indicate that there currently is a higher need for treatment. Other studies that observed German samples during the pandemic also found a nonsignificant association of the presence of chronic conditions and telemedicine use [22,23]. Moreover, necessary treatment for patients with severe diseases (eg, physical examination, cancer treatment) was potentially more likely to be in person and still available during the pandemic. Future telemedicine services might be less suitable for these patient groups. Further research is needed to gain a better understanding of the possible impact of physical illness on telemedicine service use, especially in German samples.

Since few recent studies have examined the association of psychological symptoms with telemedicine use, our findings contribute to existing knowledge concerning psychological determinants during the pandemic in Germany. We observed that symptoms of anxiety, depression, or loneliness increased the likelihood of telemedicine use in middle-aged and older adults. Similar to our results, Hajek and König [30] observed a significant positive association of loneliness and telemedicine use in middle-aged and older adults during the pandemic in Germany. Likewise, Rauschenberg et al [33] found that psychological distress was associated with the current use of mobile health apps in a representative sample of youth aged 16-25 years from the German general population. Other studies with samples from the United States also observed a positive relationship of psychological symptoms with telemedicine use during pandemic times [54,55]. Therefore, it may be the case that findings regarding higher health care use in individuals with mental illness [56-60] can be applied to the field of telemedicine and the pandemic context. These initial findings illustrate the future potential of telemedicine in the field of mental health for middle-aged and older patients, since those who experienced psychological symptoms appeared to preferentially opt for telemedicine services. Moreover, mental health problems, such as anxiety or depression, have been shown to be positively associated with fear of COVID-19 [61]. This fear could also favor increased telemedicine use due to concerns of being infected with SARS-CoV-2 during in-person health care visits—this association was also found in our sample.
Additionally, other COVID-19–related determinants that were included in our sample (ie, vaccine status, COVID-19 infection history in oneself or close others, and preventive medication), were not significantly associated with telemedicine use. This is in contrast with findings from German samples that looked at younger and adult individuals during the pandemic [23,32,33] and found significant associations of COVID-19–related factors with telemedicine use (ie, higher perceived severity of COVID-19 infection; having had COVID-19 infection; subjective COVID-19–related challenges; and COVID-19–related cognitive preoccupation, anxiety, and worries). However, when looking at a similar sample to our study, which consisted of middle-aged and older adults during the pandemic in Germany, Hajek and König [30] did not find significant associations of COVID-19–related factors with telemedicine service use. This could potentially mean that middle-aged and older individuals’ decision to use telemedicine was less influenced by COVID-19–related factors than in the general adult or younger German population.

Strengths and Limitations

This study is one of only a few studies that explore determinants of use of remote medical consultations in German middle-aged and older adults. The nationally representative sample of the widely acknowledged SHARE panel study provides insight into the largest age group of the German population. The data were collected during the COVID-19 pandemic and therefore account for the unique circumstances that individuals were faced with during that time.

However, some limitations should be considered. Telemedicine service use was measured using only one item, which indicated experience with remote medical consultations over the phone, computer, or any other electronic means since the outbreak of the pandemic. Therefore, we did not differentiate between different patient groups, telemedicine modalities, or frequency of use. This should be explored further in future studies. Furthermore, the survey covered a limited selection of socioeconomic, health, and psychosocial aspects. Future studies should include more extensive instruments and variables to make more reliable and comprehensive conclusions. In addition, the majority of our sample did not use telemedicine and case numbers were small for some of the included determinants. This lack of statistical power might explain why some of the tested relationships did not reach statistical significance. Consequently, future studies with very large German samples are needed. Furthermore, this analysis was based on self-reported cross-sectional data, and it is therefore difficult to identify causal relationships. Finally, we only focused on the German context. Future research should also consider cross-cultural differences in use and determinants of telemedicine to better understand potential barriers and facilitators in different cultural contexts and improve worldwide implementations.

Conclusions

To achieve high rates of use and widespread acceptance of telemedicine, it is essential to understand determinants of telemedicine service use in middle-aged and older individuals. Our study findings stress the link between psychological symptoms and telemedicine use in Germany during the COVID-19 pandemic. Middle-aged and older adults seem to have used telemedicine services according to psychological needs and health limitations. One may conclude that, especially when they had psychological symptoms, middle-aged and older individuals accepted telemedicine as a service option. While socioeconomic and social factors were not associated with telemedicine service use, the associations of other health- and COVID-19-related determinants with use behavior remain unclear.

Future (longitudinal) studies are therefore required to confirm these initial findings and clarify whether they also apply to postpandemic circumstances, where widespread in-person visit availability returned. Some patients might have used telemedicine only because they had no other option. However, remote consultations might be especially suited for specific patient groups or forms of treatment and will remain part of postpandemic routine care. Furthermore, use of (remote) blended therapy might increase in the postpandemic context, as it combines the strengths of remote and in-person visits and can be adapted to individual patient preferences. Moreover, potential differences in determinants of telemedicine use between different service types (eg, asynchronous vs synchronous services) or patient groups (eg, mental health vs oncology patients) should be further investigated. Finally, it remains to be explored to what extent determinants of telemedicine use differ from determinants of general health care use, which could help to identify target groups and appropriate fields of application for future telemedicine services. This could be examined in the postpandemic context where both forms of services, in-person and telemedicine visits, are likely to be available to patients.

Acknowledgments

We acknowledge financial support from the Open Access Publication Fund of UKE (Universitätsklinikum Hamburg-Eppendorf) and the Deutsche Forschungsgemeinschaft (German Research Foundation).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Results from logistic regression and Firth logistic regression for determinants of telemedicine service use during the COVID-19 pandemic.
References


Abbreviations

OR: odds ratio
SHARE: Survey of Health, Ageing and Retirement in Europe
WHO: World Health Organization

© Ariana Neumann, Hans-Helmut König, André Hajek. Originally published in JMIR Aging (https://aging.jmir.org), 23.4.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License.
Barriers to and Facilitators of Older People's Engagement With Web-Based Services: Qualitative Study of Adults Aged >75 Years

Annemarie Money1,2, BA, MSc, PhD; Alex Hall1, BA, Grad Dip, MSc, PhD; Danielle Harris1,2, BSc, MSc; Charlotte Eost-Telling1,2, BSc, MSc, PhD; Jane McDermott1, BA, MA; Chris Todd1,2,3,4, BA, MA, PhD

1School of Health Sciences, University of Manchester, Manchester, United Kingdom
2National Institute for Health and Care Research, Applied Research Collaboration Greater Manchester, Division of Nursing Midwifery and Social Work, Faculty of Biology, Medicine and Health, Manchester, United Kingdom
3Manchester Institute for Collaborative Research on Ageing, The University of Manchester, Manchester, United Kingdom
4Manchester University NHS Foundation Trust, Manchester, United Kingdom

Corresponding Author:
Annemarie Money, BA, MSc, PhD
School of Health Sciences
University of Manchester
5th Floor, Jean McFarlane Building
Manchester, M13 9PL
United Kingdom
Phone: 44 0161 306 7777
Email: annemarie.money@manchester.ac.uk

Abstract

Background: The COVID-19 pandemic has accelerated the shift toward the digital provision of many public services, including health and social care, public administration, and financial and leisure services. COVID-19 services including test appointments, results, vaccination appointments and more were primarily delivered through digital channels to the public. Many social, cultural, and economic activities (appointments, ticket bookings, tax and utility payments, shopping, etc) have transitioned to web-based platforms. To use web-based public services, individuals must be digitally included. This is influenced by 3 main factors: access (whether individuals have access to the internet), ability (having the requisite skills and confidence to participate over the web), and affordability (ability to pay for infrastructure [equipment] and data packages). Many older adults, especially those aged >75 years, are still digitally excluded.

Objective: This study aims to explore the views of adults aged >75 years on accessing public services digitally.

Methods: We conducted semistructured qualitative interviews with a variety of adults aged ≥75 years residing in Greater Manchester, United Kingdom. We also interviewed community support workers. Thematic analysis was used to identify the key themes from the data.

Results: Overall, 24 older adults (mean age 81, SD 4.54 y; 14/24, 58% female; 23/24, 96% White British; and 18/24, 75% digitally engaged to some extent) and 2 support workers participated. A total of five themes were identified as key in understanding issues around motivation, engagement, and participation: (1) “initial motivation to participate digitally”—for example, maintaining social connections and gaining skills to be able to connect with family and friends; (2) “narrow use and restricted activity on the web”—undertaking limited tasks on the web and in a modified manner, for example, limited use of web-based public services and selected use of specific services, such as checking but never transferring funds during web-based banking; (3) “impact of digital participation on well-being”—choosing to go to the shops or general practitioner’s surgery to get out of the house and get some exercise; (4) “the last generation?”—respondents feeling that there were generational barriers to adapting to new technology and change; and (5) “making digital accessible”—understanding the support needed to keep those engaged on the web.

Conclusions: As we transition toward greater digitalization of public services, it is crucial to incorporate the perspectives of older people. Failing to do so risks excluding them from accessing services they greatly rely on and need.

(JMIR Aging 2024;7:e46522) doi:10.2196/46522

KEYWORDS
digital exclusion; digital inclusion; older people; technology; aged; web-based; internet
**Introduction**

**Background**

The shift toward digital technologies to provide access to essential and nonessential public services, such as health care, public administrative services, utilities, financial services, and leisure, has been accelerated by the COVID-19 pandemic [1-3]. In many countries, including the United Kingdom (the setting for the research reported in this paper), this shift has also been driven by national policy in attempts to maintain access to essential services during periods of social lockdown [4-6]. Many social, cultural, and economic activities, such as appointment or ticket booking, have transitioned to web-based platforms [7], whereas many COVID-19-related initiatives, such as booking and recording COVID-19 tests and vaccinations, were primarily offered through digital platforms to the public. In particular, during the COVID-19 lockdown periods, being on the web became critical for maintaining social ties and combating loneliness [8-10].

This may be convenient for many people, and there is evidence to show that digital inclusion, accelerated by the pandemic, has increased in recent years [11]. However, although in the United Kingdom, the “digital divide” (the gap between those who do and those who do not have access to new forms of information technology [12]) may have narrowed in recent years, the impact has not been felt equally and has widened for some groups [13,14]. The digital divide remains especially wide for older adults aged ≥75 years; this is further exacerbated by living alone, having a limiting long-term condition, and being financially susceptible [15,16]. An analysis of the English Longitudinal Study of Ageing data conducted early in the pandemic (June and July 2020) showed that 45% of adults aged 52 to 64 years and 41% of adults aged 65 to 74 years used the internet more since the COVID-19 outbreak, but only 24% of those aged >75 increased their use and 9% were using it less [17].

Digital exclusion, in its broadest definition [15], relates to three connected aspects: (1) access, that is, whether individuals have access to the internet at home or elsewhere; (2) ability, that is, having the skills and confidence to participate over the web; and (3) affordability, that is, the ability to pay for infrastructure (equipment) and adequate data packages [16]. In the United Kingdom, recent data from the national communications regulator (Ofcom) suggests that 6% of UK households do not have home internet access, but this figure rises to 26% of people aged ≥75 years [18]. It is also acknowledged that an additional 2 million households are experiencing financial difficulty, and this will likely increase given the cost-of-living crisis from 2021 to 2023 [19,20]. Digital exclusion is viewed as a “super” social determinant of health [21] as it impacts a variety of areas of life, including leading to poorer health outcomes [22,23] and challenges with employment, housing, education, and finance. It disproportionately affects many people, including people with low incomes, people living in social housing, people living with disabilities, people in rural areas, and people for whom English is not their first language, as well as other marginalized groups. Although all these factors are important indicators of who is likely to be digitally excluded, age remains the biggest indicator.

According to data from the UK Office for National Statistics for 2020, a total of 99% of adults aged 16 to 44 years were recent internet users compared with only 54% of adults aged ≥75 years [24]. It is vital to understand the complexities of how digital exclusion exacerbates health and social inequalities so that adequate responsive action can be considered. For example, it is not simply the case that ensuring internet connectivity will mitigate digital exclusion. There is a need to understand structural challenges; financial barriers; digital literacy; and other aspects, such as the impacts of various health conditions and disabilities and concerns about privacy and data protection [25-28].

A recent scoping review explored the barriers to and facilitators of older people’s digital engagement across the spectrum of nonuse through sustained use [29]. This review found that there are substantial overlaps between barriers and facilitators; for example, lack of knowledge of digital technologies is a barrier, prior knowledge is a facilitator, perceived lack of personal capability is a barrier, and a positive attitude toward oneself is a facilitator. The review also found a substantial gap regarding the determinants of technological nonuse. Although this review provides a thorough scope of the literature, it included studies involving participants with a mean age ≥26 years and did not offer any further stratification by age. There is a need for a more nuanced focus on older groups of older people. Current insights into digital technology use often exclude a specific focus on people aged ≥75 years. The coverage of age groups in several major reports and surveys often stops at the age of 74 years or includes all people aged ≥65 years as 1 homogeneous age category [30].

**Objectives**

Given the increased risk of digital exclusion among older age groups and the fact that this may have been exacerbated by the COVID-19 pandemic, it is vital to gain a deeper understanding of the use, attitudes, and preferences of people aged ≥75 years.

In the United Kingdom, this gap has been recognized as a policy priority. In Greater Manchester, a city-region in North West England with a population of 2,867,800 in 2021, as many as 1.2 million residents are estimated to be limited digital users because of exclusion or personal preferences, with a substantial proportion of these being people in later life. To address this digital divide, the Greater Manchester Combined Authority (GMCA) established a Digital Inclusion Action Network and Taskforce in October 2020 [31], with the ambition to make Greater Manchester a 100% digitally enabled UK city-region. Older adults aged ≥75 years were included as a critical population group for targeted action related to digital inclusion. The aim of the paper is to report findings from a qualitative study exploring the views of adults aged ≥75 years on accessing web-based public services.

**Methods**

**Study Design**

The National Institute for Health and Care Research Applied Research Collaboration Greater Manchester was commissioned by the GMCA to gather insight into the barriers to and
facilitators of older adults’ (>75 y) digital participation within the region. To address this, we conducted semistructured qualitative interviews with adults aged ≥75 years residing in Greater Manchester.

**Sampling, Recruitment, and Data Collection**

Purposive and convenience sampling were used to identify and recruit participants. Adapting to web-based rather than face-to-face data collection during the COVID-19 pandemic lockdown meant that we had to use an approach that allowed us to use our existing networks and recruit via several third-party organizations affiliated with the GMCA Ageing Hub. We recruited a range of older adults; these included older adults who were fully engaged and participating in many web-based activities (often enrolled in a local support program), those who were just starting to receive support to get on the web, those who were previously engaged but were now lapsed users, and those who had no interest in getting on the web or using a computer or device at all. We also recruited community support workers to learn from the approaches they used to continue to engage with their communities, particularly during the COVID-19 pandemic period (from 2020 to 2022) [32].

Data were collected via semistructured interviews (conducted by AM) with a topic guide (Multimedia Appendix 1) developed from a rapid review of the literature [33] and from input from the project oversight team.

**Data Analysis**

Interviews were audio recorded, transcribed, and exported to NVivo Pro (version 12) software for data management [34]. Using a thematic analysis approach [35], initial themes were identified from the transcripts and indexed to develop the analytical categories. Via a process of constant comparison [36], these categories were reviewed and refined by 2 researchers (AM and DH), and any ambiguities in the coding framework were reconciled by a thorough discussion with the research team. All interviews were then fully coded using NVivo Pro for qualitative analysis (AM and DH).

**Ethical Considerations**

Ethics approval was granted by the University of Manchester Proportionate Research Ethics Committee (2021-12638-20811). All interviews were conducted virtually (by telephone or other remote means agreed upon with the participants) at a time convenient to the participants; the participants provided informed consent before data collection. The data were collected between October 2021 and February 2022. All data provided was anonymised by the research team with any personal identifying information removed. All participants received a £15 'Love2Shop' voucher as a thank you for their participation in the interview.

**Results**

**Overview**

The final sample comprised 26 interviews: 92% (24/26) with older adults aged >75 years and 8% (2/26) with community digital support officers. The older adult sample had a mean age of 81 (SD 4.54; range 75-91) years; 58% (14/24) of the participants were female, 96% (23/24) were White, and 4% (1/24) were of South Asian background. Participants were sampled from 4 (40%) of the 10 local authority areas in Greater Manchester, 75% (18/24) of the participants were users of the internet (to some degree), and interviews lasted on average 23 (SD 8.51; range 8-60) minutes.

A total of 5 themes were identified as being key in understanding the barriers to and facilitators of motivation, engagement, and participation in using web-based public services. The key themes and subthemes are presented in Textbox 1 and discussed in detail in the subsequent sections.

**Textbox 1. Key themes and subthemes identified via thematic analysis.**

- **Theme 1: initial motivation to participate digitally**
- **Theme 2: narrow use and restricted activity on the web**
  - Preference for choice of access
  - Narrow use driven by fear
  - Lack of interest in learning new digital skills and tasks
- **Theme 3: impact of digital participation on well-being**
- **Theme 4: the “last generation”?**
- **Theme 5: making digital accessible**
  - One-to-one support
  - Mitigating physical impairments

**Theme 1: Initial Motivation to Participate Digitally**

Among older adults who were using the internet, it was evident that their decision to do so was often driven by a particular, recent need that motivated them to go on the web. Unsurprisingly, given that COVID-19 pandemic lockdowns reduced face-to-face social contact [32], one of the major motivators related to maintaining social connections and gaining skills to be able to connect with family and friends in other parts of the world. A participant stated the following:
Well, I think that’s vital really [being online]. It’s kept me alive in that, you know, I feel as if I’m speaking to people. It’s company there. I’m never isolated because I can always get in touch with somebody. So, to me, it has literally been a lifeline. [Participant 19, female, aged 82 years]

There was also the need and convenience of being able to access certain services during the lockdown, in particular web-based ordering of prescriptions and shopping. A participant said the following:

That’s the most brilliant thing I’ve ever used, Amazon. You don’t even have to go outside the door. [Participant 13, female, aged 75 years]

**Theme 2: Narrow Use and Restricted Activity on the Web**

**Preference for Choice of Access**

Although three-quarters (18/24, 75%) of the older adults interviewed were digitally engaged (to some extent), many of them were “narrow” users [11,15], in that they participated only in a handful of web-based activities or tasks. The participants reported very little interaction with web-based public administrative services (eg, local and national government services such as disabled parking applications, passport applications, driving licenses, and benefit applications). The participants were more positive about some aspects of web-based health services, particularly ordering prescriptions. When asked why they preferred to order prescriptions over the web, some commented that the system is “straightforward to use” and that they “find it very useful” to be able to order over the web, with some noting that if they did not do it over the web, it would involve them going to the surgery, “which is a bus ride away.” However, when asked about the prospect of more public services moving to web-based access, most participants—both those who used the internet and those who did not—were in consensus that digital should not be the only option provided by organizations to access a service, for a variety of reasons. A participant stated the following:

The jabs that I had, it was telling me to go online, that’s an example, and I phoned up my doctor’s surgery and said, look, I can’t go online, right, so they did the appointment for me. And also I had a bit of an argument with [large retail pharmacist named] and other stores like [pharmacy chain] because the flow thing, you know the flow thing, [Lateral Flow Test—rapid antigen test for COVID-19] you’ve got to go online although you didn’t get any because they were sold out, they were out of them all the time...it’s just ridiculous, nobody thinks about the older people. [Participant 14, male, aged 83 years]

Another participant said the following:

Personally I don’t think you can beat seeing the doctor face to face, they can pick up on your body language, colour of your skin. I think there’s lots of things that you can pick up on face to face. So, I do think seeing the doctor face to face is essential for the majority, I really do. I think going online for some things is good, but I do think if you’re not very sure about what you’re doing, I think...I would imagine it could cause a lot of stress, if there’s no alternative... [Participant 24, female, aged 76 years]

**Narrow Use Driven by Fear**

Many participants adopted a granular approach to use, in which they had specific and limited web-based tasks they would undertake within particular domains of activity, such as banking or shopping. For example, many were happy to log on to banking apps or websites and view their balance—that is, to monitor their account—but stopped short of undertaking any transactions. Reasons for this limited use included a concern about having personal details “out there,” pressing the wrong button and sending the money to the wrong place, and a fear of being scammed. A participant stated the following:

...No, I won’t do finance at all, PayPal or anything. I really don’t trust it because there are so many scams around, erm, I just think it’s too easy, if you press the wrong button and its gone to Timbuktu, no I definitely won’t have anything to do with online banking, and it’s a shame because I know that I would shop online and it would save a trip to the Post Office or the bank or whatever, but I just wouldn’t trust it... [Participant 7, female, aged 79 years]

Another participant stated the following:

...I know I’ve got the banking online on the tablet if I want to use it for transactions, but I’m just quite happy seeing what I’ve got at the moment. I don’t really feel confident enough to do transactions. I always worry, God if I do something wrong, I’m in trouble. [Participant 24, female, aged 76 years]

In addition, another participant said the following:

I just feel as though I don’t want to be divulging too much information about myself to the wide world, if you know what I mean. [Participant 11, female, aged 87 years]

Community digital support officers highlighted fear and concerns around the safe sharing of personal information as a key barrier to engagement among older adults. In particular, media reports highlighting scams and frauds were deemed to exacerbate this barrier. The community officers were aware of the need to inform people of the potential risks, but “more positive campaigns about [the benefits of] using the internet [for older people] are needed” (digital support officer 2), as the negative stories reported on television were seen to deter older adults from benefiting from available web-based services. They reported that the word “scam” really “puts fear into older persons” and the media “cherry pick” the very worst scams to the point that people are convinced that these are happening on their very doorsteps:

...And this is no joke, I have had people say to me that they think there are people outside their house on a laptop in their car, you know they are parked on the
street trying to use their Wi fi to scam them. That sort of thing. [Digital support officer 1]

Lack of Interest in Learning New Digital Skills and Tasks

Where participants did use web-based services, once their initial needs had been met and they had gained the necessary skills to complete an activity, many lacked interest or were reluctant to undertake additional tasks or learn new activities. They were happy to maintain the skills and knowledge gained to undertake the tasks or activities that initially prompted them to get on the web. A participant stated the following:

I wouldn’t use it for much really. I’m not ambitious about it. I have done what I wanted to do and anything else that I gained, it’s a bonus. [Participant 10, female, aged 91 years]

Another participant stated the following:

Are there any tasks that you haven’t yet done online that you think you might want to try or you want to do in the future? [Interviewer]

No, because I can use the computer and I can use the phone and the tablet for anything that I personally need to do. [Participant 20, male, aged 76 years]

In addition, another participant said the following:

And are there things that you would want to do, that maybe you don’t yet know how to do, or you’d need some support to be shown how to do it? [Interviewer]

I don’t think so. I think I do what I need to do... [Participant 13, female, aged 75 years]

Theme 3: Impact of Digital Participation on Well-Being

For some participants, who were not digitally engaged and had no interest in getting on the web, social connections and social interaction were cited as a reason for not engaging. One participant stated the following:

No, I just think I’ve never been interested. I feel that if I did use something I’d be on my own doing it, and I don’t like being on my own. Years ago I had one of these knitting machines and I had it for a while, and I hated it because it meant I was sat on my own knitting, and I don’t want to do that. I like to go out and meet people while I can. [Participant 3, female, aged 90 years]

Another participant stated the following:

This is the problem, lots of people don’t talk anymore. They know...they don’t know any other way of corresponding, getting in touch with people. I mean they go on the internet. They text, they don’t talk...I mean you go out for a walk and you can more or less guarantee at least 50 per cent of people walk around with their phone. [Participant 5, male, aged 83 years]

The participants also spoke about digital engagement in relation to aspects of physical and mental well-being. For example, for some, not using web-based services, such as to make a general practitioner (family physician in the United Kingdom) appointment, was seen as a positive because they had to get “out of the house” and, in doing so, had the benefit of getting “a little bit of exercise.” For others, there was the acknowledgment of the advantages of being able to shop over the web during lockdowns, but now that restrictions had lifted, they had reverted to their preference to shop in person, which again was seen as an opportunity for exercise. Others talked about how it was sometimes “too easy” to depend on the internet to find out information that they could not immediately call to mind, and this was spoken about in terms of brain health and keeping the mind active. A participant stated the following:

So, in terms of doctors’ appointments and things, you can still get to the surgery or you could ring. Is that something that you prefer to do? [Interviewer]

Yes, I can ring on my landline and talk to a receptionist, or just toddle myself down to the surgery and go face to face with them, you know...Not that I would get an appointment any quicker with the doctor but, you know...And it gives me a little bit of exercise. [Participant 18, female, aged 79 years]

Another participant stated:

I’ve gone back to going out because you get a little bit of exercise, you know. So yeah, I don’t shop for groceries online anymore, no, I always go to the shop. [Participant 20, male, aged 76 years]

Another participant stated the following:

...A couple of days ago, it sounds ridiculous this...I thought, what’s the name of that pub at the top of Lancashire Road? I mean, what the hell I thought about that for, and I could not for the life of me, and I thought, no...remember it, because you do know it. And this morning, it’s come to me, The Hinds Head it’s called. [Participant 13, female, aged 75 years]

Okay, so you resisted the urge to find out? [Interviewer]

I did, yes, I did. I thought, no, that’s too easy. [Participant 13, female, aged 75 years]

Theme 4: The “Last Generation”?

Many participants, particularly those who were not on the web, spoke about barriers regarding generational issues and how they felt they might be the “last generation” to experience difficulty with digital participation:

But I do think that we’re the last generation, almost the last generation that this will affect. Because from being babies now they have iPads now and what-have-you, don’t they? It’s just second nature to them. It puts you to shame when you watch them. [Participant 12, female, aged 81 years]

This often went hand-in-hand with the perception that these difficulties were unique to their generation and that the younger generations experienced little difficulty in adapting to or embracing new technology. A participant stated the following:

Well, do you know what, to be honest with you, I could say I’m at the end of a generation. Because if you
think of the youngsters today now and you think of...like my sons have no problem with this, that’s another generation, and then the one below that is the youngsters, yeah, this will never happen again...it’s unlucky, I’m at the end of a generation. [Participant 14, male, aged 83 years]

Another participant stated the following:

The youngsters, from school onward, they know nothing but the internet. So everything is being geared toward them. And we older people, in my generation, have had to start learning various things which become harder and harder. It’s second nature to younger people, to the 30s, 40s. [Participant 5, male, aged 83 years]

Theme 5: Making Digital Accessible

For those older adults who were participating (to some extent) in web-based activities, it was important to discuss barriers and facilitators that might require consideration to keep older adults engaged and supported.

One-to-One Support

Participants emphasized the importance of having patient assistance while navigating tasks on the web. They also highlighted the value of receiving written instructions and having tasks demonstrated multiple times. The participants also valued the one-to-one support given to them but stressed that this needed to be ongoing support, noting that sometimes they would “get the hang of” one task (eg, shopping) only to find that the next time they logged on to the website, the landing page may have changed, which would “throw them off” and result in them feeling unsure whether they could continue in the manner they had been shown. A participant stated the following:

I’d love someone to sit and show me so I can write it down and if I get stuck I know how to do it myself. [Participant 19, female, aged 82 years]

Other participants stated the following:

She writes things down for me, because I can’t always remember what I’ve been told. If I’ve got it there in black and white then I can follow it. It does help. [Participant 2, female, aged 76 years]

Yeah, sometimes it doesn’t click immediately and you need them to go over it again. So you need somebody who’s got a little bit of patience. [Participant 21, male, aged 78 years]

I also like the fact that I can have it on my knee in the lounge or the chair that I’m in. [Participant 19, female, aged 82 years]

Discussion

Principal Findings

The aim of this project was to explore the views of adults aged ≥75 years on accessing web-based public services in response to a policy initiative to further understand older people’s digital behavior and engagement. The analysis of the semistructured interviews identified 5 themes that were key to understanding some of the barriers and facilitators experienced by the older adults participating in this project. The key facilitators included responses to meeting certain needs (particularly during national lockdowns) such as food shopping, ordering prescriptions, and staying connected with family and loved ones. The identified barriers included fear of scams and misuse of personal information, lack of ongoing support to maintain or learn new skills, preference for face-to-face interactions (especially for health appointments), and a wider generational belief (held by many) that difficulties getting on the web were “unique” to their generation and that older adults found it difficult to adapt or embrace new technology. Crucially, we also found that the potential unintended consequences of the benefits offered by digital technologies to access public services could be seen as a barrier to their use. In particular, this included their ease of access to information and their convenience, which were seen to reduce the need for people to engage cognitively elsewhere or to leave the house, thereby denying them exercise and social interaction opportunities.

Theoretically, there are several models that attempt to explain digital engagement and uptake. Two of the most well-known and widely used are the Technology Acceptance Model [37] and the Unified Theory of Acceptance and Use of Technology [38]. However, these models are primarily used to quantify the acceptance of technologies rather than to provide qualitative
insights [29]. One straightforward categorization to facilitate an understanding of different “types” of older adults’ digital behavior suggests they may fall into 1 of 4 groups—“engaged,” “disheartened,” “transitional,” or “uninterested” [39]. “Engaged” refers to those older adults who believe they are capable of learning and perceive a value in using digital technology, that is, believe the internet is useful to them personally. “Disheartened” users also believe the internet to be useful and usually have more need for digital services but are worried about safety and associated risks and feel less confident in their ability and skills. “Transitional” older adults often have the highest need for use of digital services, but they are frequently lapsed users with narrow, if any, experience of digital engagement other than for social media purposes. “Uninterested” older adults do not perceive value in web-based activity and often have others access websites on their behalf. They usually have strong social connections and can be resistant to using the web. Although these categories are helpful in starting to think about digital behavior and potentially offer insights into how to support different “types” of older adults, the interviews presented here show that older adults’ digital behavior may not always be neatly classified into 1 type of user versus another. A large proportion of the respondents could be classified as digitally “engaged” in the sense that they were keen to go on the web, felt capable of learning, and had many of the skills deemed “essential” [40] for digital participation.

However, it was not possible to map these older adults to 1 “type” as there was often an overlapping of categories (particularly “engaged” and “disheartened”), which required a more nuanced understanding of what digital participation means for older adults. During discussions around the motivations behind getting on the web for our older adults (particularly during the pandemic), the initial “engagement” was evident; however, this engagement for many appeared to ebb away, and subsequently, many of them fitted the descriptions for other categories, for example, “disheartened” users. Although many of our older adults admitted a perceived value in accessing the internet and participating digitally (ie, “engaged”), a lack of confidence, lack of support, or fear of sharing information would often result in them becoming “disheartened,” disengaging from aspects of internet use, and not taking full advantage of the services available to them (eg, restricted use of web-based banking). Understanding that many older adults will not “fit” into 1 category highlights the need for a more individualized and nuanced approach to tailoring digital support services [41].

When considering the wider impacts of digital technology transformation, older adults’ limited use of web-based activities such as banking and concerns over data protection were also amplified by a lack of confidence and skills. Added to this were the needs of older people to get out and to socialize, with health-related appointments, shopping, and banking forming part of how participants stayed active and well in their communities. Work undertaken with a range of individuals, including older adults, during the pandemic found that a move to internet-based general practitioner and health appointments was sometimes problematic for this group for a variety of reasons, including a lack of skills and confidence, no interest in engaging on the web, and a lack of trust [42-44]. In addition, web-based platforms provide a very limited 2D view of a person and their circumstances [45,46]. As such, it is critical that people are encouraged to leave their homes and are able to access and attend face-to-face appointments. We already know that far too many older people are sedentary and do not achieve the recommended levels of physical activity [47,48]. This worsened during the COVID-19 pandemic, negatively impacting health [49]. Therefore, adding to this burden by substituting physical activity with digital engagement should be avoided.

The pandemic and its successive lockdowns have moved much public and social activity to web-based platforms. Digital exclusion is often discussed in terms of “hard” (eg, never having used the internet or having no internet access) or “soft” (eg, improving digital engagement, skill level, or confidence) [7]. These changes driven by the pandemic are said to have improved “hard” exclusion for the general population. However, in terms of improving the “softer” areas of exclusion, the pandemic has done little to close the digital divide, particularly for older adults [50]. There is evidence to show rates of internet use increasing faster among younger cohorts and declining among older cohorts, demonstrating the digital divide naturally closing in time as generations who experience high levels of digital exclusion are replaced by younger generations who embrace and adapt to technological change [51,52]. This idea came through strongly in the interviews conducted, with many older adults expressing this view. However, it was not clear whether they attributed this to the impact of the rapid digitalization brought about by the pandemic or to the impact of a more general move toward digitalization over a longer period. However, the rapid development of technology combined with an individual’s changes in physical health has been shown to worsen feelings of being unable to keep up or feeling too old to embrace new technology among older adults [53]. There is also the view that older adults can often internalize agism [54] and accede to the stereotype that they are not able to master technology and so do not attempt it. It seems plausible that the rapid increase in digitalization during the pandemic may have exacerbated these feelings of being left behind, but it is also important to note that inaccurate perceptions of young people as fluent technology users may be driven by a broad range of factors, including media representation, agism, and other social constructs related to digital inclusion and exclusion. Although the impact of the 2020 to 2023 COVID-19 pandemic brought many of these issues to the forefront in discussions around digital participation and the impact of the divide for older adults, these are not new issues related only to older adults’ experiences during the pandemic [50].

This study highlights several digital technology features that have delivered positive outcomes for people aged ≥75 years. Among those we interviewed, there was a preference toward using tablets [55] as well as a willingness to order prescriptions and engage in web-based shopping. Shared learning across public institutions on aspects of digital technology transformation that have been delivered successfully for adults aged >75 years would be beneficial. For example, what can we learn from the experience of web-based ordering of prescriptions that could inform other public service digitalization infrastructure and processes? There is also the importance of...
ensuring digitalized services are fully accessible to all, including those living with long-term health conditions that may impact their ability to use digital devices (eg, arthritis) or access content easily (websites, text, etc). For older adults who wish to participate on the web, building confidence in undertaking tasks, such as banking, via support that is task focused and repeated is crucial [56,57].

**Strengths and Weaknesses**

Current data and insight into digital technology use may exclude people aged ≥75 years or may lack a specific focus on this age group. Often, available data on older adults’ use stops around the age of 74 years or it provides information on all individuals aged ≥65 years. Given the increased risk of digital exclusion among older age groups and the fact that this may have been exacerbated by the COVID-19 pandemic, it is vital to gain a deeper understanding of the use, attitudes, and preferences of people aged >75 years. Although this research was relatively small in number, its strength is that it focuses on those aged >75 years (average respondent aged 81 years). A key message to take away from the work is one of choice: that those aged ≥75 are not digitally homogenous but rather require a range of options, both digital and nondigital, that will enable them to engage in ways that work best for them and do not further exacerbate digital inequalities [23,58-60]. A key finding adding to the literature [29] is that the benefits of technologies, such as ease of access to information and convenience, may actually have unintended consequences that put older adults off using them. This includes a desire among some participants to continue to access some public services in person for the indirect benefit of physical activity while doing so. This finding is particularly important in light of other healthy aging policies that promote physical activity to improve disability-free life expectancy [61].

Future research should investigate the experiences of older adults from ethnic minority groups. Although this study aimed to be as inclusive as possible, the recruitment of older adults from diverse ethnic minority groups proved challenging. It would also be beneficial to examine the influence of age-related sensory changes on digital inclusion. Investigating the preferences and experiences of older adults with hearing or visual impairments would provide valuable insights. Physical distancing and stay-at-home restrictions during pandemic lockdowns meant that our recruitment strategy had to be adapted to make use of our existing networks and ties with third-party organizations to be able to recruit older adults for the study. A reliance on web-based means of recruitment resulted in a more digitally engaged sample of older adults being recruited than originally intended, although the levels of engagement varied among the older adults. In an ideal setting, a study of this nature would have been conducted with older adults in a face-to-face setting rather than via telephone or video interviews [62-64]. However, data collection during the COVID-19 pandemic meant this approach was not possible, and this will have had an impact on the final sample of older adults, with a larger proportion of adults who were digitally engaged taking part.

**Conclusions**

The shift to digital delivery of public services, both throughout the pandemic and more generally as a driving force for future service provision, requires a focus on the needs and preferences of older people so that they are not excluded from service access. Mitigation against digital exclusion is a core component of 1 of the strategic priorities to reduce health inequalities across England [52]. It is vital that the needs and preferences of people of all ages are considered, particularly those aged >75 years, who are often underrepresented in research. Attempts to classify “types” of digital users may be a useful heuristic for thinking about digital engagement, but the boundaries between categories are permeable and complex. Those aged ≥75 years are not a digitally homogenous group but rather require a range of options, both digital and nondigital, that will enable them to access services without further exacerbating digital inequalities.

**Acknowledgments**

This project was funded by National Health Service (NHS) England through the Greater Manchester Health and Social Care Partnership and supported by the National Institute for Health and Care Research (NIHR) Applied Research Collaboration Greater Manchester (NIHR200174) and an NIHR Senior Investigator Award to CT (NIHR200299). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or its partner organizations. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results. The authors would like to thank all participants for their valuable insights and time in taking part in the interviews. The authors thank Alison Littlewood for her enthusiastic program administration and the Greater Manchester Aging Hub at the Greater Manchester Combined Authority for their support throughout this study.

**Authors’ Contributions**

The conceptualization and design of the study was undertaken by AM, JM, and AH. Data collection was undertaken by AM. AM and DH coded and analyzed the data. AM drafted the original paper as first author and DH, JM, AH, CET, and CT worked on reviewing and editing subsequent drafts. All the authors have read and agreed to the published version of the manuscript.

**Conflicts of Interest**

None declared.

Multimedia Appendix 1
References


39. I am connected: new approaches to supporting people in later life online. Good Things Foundation. URL: https://www.goodthingsfoundation.org/insights/i-am-connected/ [accessed 2024-02-16]


42. Tewary S, Cook N, Dezine M, Shnayder O, Pandya N. Supporting vulnerable older adults with telehealth through wellness calls and tablet distribution during COVID-19: quality improvement project. JMIR Form Res 2023 Sep 11;7:e34392 [FREE Full text] [doi: 10.2196/34392] [Medline: 37695647]


48. Skelton DA, Mavroeidi A. How do muscle and bone strengthening and balance activities (MBSBA) vary across the life course, and are there particular ages where MBSBA are most important? JFSF 2018 Jun;03(02):74-84. [doi: 10.22540/jfsf.03.074]


Abbreviations

GMCA: Greater Manchester Combined Authority
Online Cognitive Stimulation Therapy for Dementia in Brazil and India: Acceptability, Feasibility, and Lessons for Implementation

Emily Fisher1, MSc; Shreenila Venkatesan2, MSc; Pedro Benevides3, MSc; Elodie Bertrand4, PhD; Paula Schimidt Brum5, PhD; Céline El Baou1, MSc; Cleusa P Ferri5, PhD; Jane Fossey6, PhD; Maria Jelen1, BSc; Jerson Laks3, MD, PhD; Lisa Liu1, BSc; Daniel C Mograbi3, PhD; Nirupama Natarajan3, MScPH; Renata Naylor3, MSc; Despina Pantouli1, BSc; Vaishnavi Ramanujam2, MD; Thara Rangaswamy2, MD, PhD; Raquel L Santos de Carvalho3,9, PhD; Charlotte Stoner10, PhD; Sridhar Vaitheswaran3, MD; Aimee Spector1, DClinPsy, PhD

1University College London, London, United Kingdom
2Dementia Care in Schizophrenia Research Foundation, Chennai, India
3Pontifical Catholic University of Rio de Janeiro, Rio de Janeiro, Brazil
4Université Paris Cité, Paris, France
5Universidade Federal de Sao Paulo, Sao Paulo, Brazil
6University of Exeter, Exeter, United Kingdom
7Federal University of Rio de Janeiro, Rio de Janeiro, Brazil
8University of Pittsburgh, Pittsburgh, PA, United States
9Universidade do Grande Rio, Rio de Janeiro, Brazil
10University of Greenwich, London, United Kingdom

Corresponding Author:
Emily Fisher, MSc
University College London
Gower Street
London, WC1E 6BT
United Kingdom
Phone: 44 20 7679 5770
Email: emily.fisher@ucl.ac.uk

Abstract

Background: Cognitive stimulation therapy (CST) is an evidence-based, group psychosocial intervention for people with dementia, and it has a positive impact on cognition and quality of life. CST has been culturally adapted for use globally. It was developed as a face-to-face intervention but has recently been adapted for online delivery.

Objective: In this study, we aimed to explore the feasibility and acceptability of online or virtual CST (vCST) delivery in India and Brazil, emphasizing barriers and facilitators to implementation.

Methods: A single-group, multisite, mixed methods, feasibility study was conducted, with nested qualitative interviews. Primary feasibility outcomes were recruitment rate, attendance, attrition, acceptability, and outcome measure completion. Exploratory pre- and postintervention measures, including cognition and quality of life, were assessed. Qualitative interviews were conducted with people with dementia, family caregivers, and group and organizational leaders following intervention delivery, and the data were analyzed using the Consolidated Framework for Implementation Research.

Results: A total of 17 vCST group sessions with 59 participants were conducted for 7 weeks, with 53% (31/59) of participants attending all 14 sessions. Attrition rate was 7% (4/59), and outcome measure completion rate at follow-up was 68% (40/59). Interviews took place with 36 stakeholders. vCST was acceptable to participants and group leaders and enabled vital access to services during pandemic restrictions. While online services broadened geographic access, challenges emerged concerning inadequate computer literacy, poor technology access, and establishing interpersonal connections online. Exploratory, uncontrolled analyses indicated positive trends in quality of life but negative trends in cognition and activities of daily living, but these results were not statistically significant.
Conclusions: vCST demonstrated feasibility and acceptability, serving as a crucial resource during the pandemic but raised challenges related to technology access, computer literacy, and long-term implementation. The study highlights the potential of vCST while emphasizing ongoing development and solutions to address implementation challenges.

doi:10.2196/55557

KEYWORDS
psychosocial; intervention; technology; COVID-19; LMIC; low and middle income countries

Introduction

Background
Dementia affects more than 57.4 million people worldwide [1]. People with dementia in low- and middle-income countries (LMICs) make up 60% of all global cases [2]; however, high-income countries account for around 74% of global expenditure on dementia [3]. There are an estimated 1.8 million people >60 years of age with dementia in Brazil [4] and 8.8 million in India, representing 5.8% and 7.4% of people >60 years of age, respectively [4-6]. While demographic and socioeconomic factors differ between Brazil and India, both countries experience low diagnosis rates, limited access to specialist treatment and care, high levels of stigma, and a lack of dementia awareness [7-9].

Worldwide, pharmacological treatments for dementia remain limited, so nonpharmacological interventions are needed to address cognitive and behavioral symptoms of dementia and improve quality of life for people with dementia and their families [10]. Many interventions are developed, but only a few are delivered at large scale and have been adopted in routine practice [11].

Cognitive stimulation therapy (CST) is a brief, manualized group program which has been found to improve cognition and quality of life in people with mild to moderate dementia [12]. It comprises themed activities that stimulate and engage participants in a social group environment led by a trained facilitator through tasks such as physical activity, word association, and discussion of current affairs [13]. Despite evidence for its effectiveness [12], cost-effectiveness [14], and its successful cultural adaptation internationally in more than 35 countries [15], CST is yet to be implemented in routine practice outside of the United Kingdom [16,17]. The 2022 World Alzheimer’s Report recommended further research and implementation of CST globally [18].

This study was conducted as part of the CST-International research program, which explored the implementation of CST in 3 LMICs [19]: Brazil, an upper middle-income country; (2) India, a lower middle-income country; and (3) Tanzania, a low-income country [20]. In each country, research teams had previously translated and culturally adapted CST and carried out feasibility studies of face-to-face CST [16,17].

CST was developed as a face-to-face intervention. However, during the CST-International study, access to face-to-face health care services was rapidly restricted due to the COVID-19 pandemic [21]. CST started to be delivered virtually in international settings including the United Kingdom and New Zealand [22,23], but a standardized protocol was not available.

An international collaboration resulted in a framework for global delivery, which was field-tested in Brazil and India, alongside the United Kingdom, Hong Kong, and Ireland [24].

Objectives
This study aimed to build upon the initial development and field-testing of the online or virtual CST (vCST) protocol in Brazil and India, with a focus on feasibility, acceptability, and implementation. Therefore, the aims of this study were to (1) explore the acceptability of vCST to people with dementia, family caregivers, CST facilitators, and service managers in Brazil and India; (2) test the feasibility of online recruitment, delivery, and assessment for CST in Brazil and India; and (3) explore factors affecting the implementation of vCST.

Methods

Design
This study was a single-group, multisite, mixed methods, feasibility study, with pre- and postintervention outcome measures followed by semistructured interviews with people with dementia, caregivers, and intervention group leaders.

Methodological Framework
This study is guided by the Consolidated Framework for Implementation Research (CFIR) [25]. The CFIR is a determinant framework, which incorporates domains that are hypothesized or found to influence implementation outcomes and has been applied in LMIC settings [26]. The five domains relate to (1) intervention characteristics (eg, evidence strength and quality as viewed by stakeholders and its core and adaptable components); (2) outer setting (eg, local attitudes to the condition of interest as well as external partnerships and financing); (3) inner setting (eg, available organizational resources and staffing); (4) characteristics of individuals involved in implementation, their need for the intervention and their capability, availability, and motivation to be involved (based on the capability, opportunity, and motivation behavior model) [27]; and (5) process of implementation (eg, assessing needs of intervention recipients, planning, and tailoring strategies). Use of the CFIR as a deductive qualitative framework enables comparisons of barriers and facilitators in other settings and for other interventions.

Participants
People with mild to moderate dementia, supported by their family caregivers, took part in vCST group sessions. For the qualitative component, people with dementia, caregivers, group leaders, and organizational decision makers were invited to participate.
participate in qualitative interviews following the completion of the vCST groups.

Setting
In Brazil, the study site was a psychology department at a university in Rio de Janeiro. In India, the study site was a mental health nongovernmental organization (NGO) in Chennai, offering outpatient, inpatient, and day center services for people with dementia. Both sites had previously been involved in cultural and virtual adaptation of CST [16,17,24] and therefore already had access to face-to-face CST manuals, vCST guidance, and CST trainers and facilitators.

Recruitment
In Brazil, recruitment took place through partnerships with memory clinics and NGOs, advertisements on social media and local media, and snowball sampling. In India, people with dementia were recruited from the patient caseload at the NGO, with additional promotion through caregiver support groups and mobile-messaging groups.

The inclusion criteria for people with dementia at both sites were as follows: they must (1) meet the ICD-10 criteria for dementia as assessed by a trained clinician [28], (2) be rated as having mild to moderate dementia on the Clinical Dementia Rating Scale [29], (3) have sufficient hearing and vision to follow conversation and comment on visual material, and (4) have the ability to participate in a online group for 1 hour.

Intervention Procedure
People with dementia were allocated to a vCST group. The vCST intervention was delivered according to the culturally adapted CST Brazil and India manuals [16,17], which had been further adapted according to the recently developed protocol for online delivery of CST [24]. Groups took place twice weekly over 7 weeks via Zoom videoconferencing software (Zoom Video Communications) between February 2021 and September 2022. Each group consisted of 3 to 5 participants. In Brazil, participants needed to use their own devices, whereas in India, devices were loaned from the NGO if needed. In Brazil, the language of instruction was Brazilian Portuguese, and in India, it was Tamil or English. Group facilitators were trained by site leaders from Brazil and India (n=7) and organizational decision makers (n=2, India only) was invited to take part in one-to-one interviews. A convenience sample of group dementia and caregivers from the first 2 vCST groups in Brazil were included in the study. Interviews were conducted with stakeholders. People with dementia and caregivers from the first 2 vCST groups in Brazil (12 dyads) and all vCST groups in India (15 dyads) were invited to take part in dyadic interviews. A convenience sample of group leaders and supporting people with dementia with additional promotion through caregiver support groups and mobile-messaging groups.

Feasibility Outcomes
The following prespecified main outcomes relating to intervention acceptability were adapted from the study by Proctor et al [30]:

- Recruitment rate: the recruitment target of 50 participants in Brazil and 15 in Chennai was intended to be achieved by September 2022. This target was calculated pragmatically based on available time and resources in each site and was deemed suitable to run enough vCST groups to explore feasibility, acceptability, and implementation issues.
- Attendance: overall attendance rate of >60%, based on the international team’s judgment and experience of running CST groups and supporting people with dementia in each setting.

- Attrition: retention rate of at least 75% of participants to the follow-up, again based on the team’s previous experience and judgment.
- Acceptability of intervention: this will be assessed through semistructured interviews (see the Semistructured Interviews section).
- Outcome measure completion: the inclusion of cognition, quality of life, activities of daily living, and caregiver burden measures are in line with those used in previous trials and studies of CST [12]. The main goal of the outcomes was to assess the feasibility and acceptability of collecting these data.

Outcome measures were translated, back-translated, and finalized by bilingual committee review. The following pre- and postintervention measures were completed by people with dementia:

- The Alzheimer’s Disease Assessment Scale-Cognitive Subscale, an internationally used 21-item measure of cognitive function [31].
- The World Health Organization Quality of Life Brief Version, a 26-item quality-of-life measure addressing 4 domains: physical health, psychological health, social relationships, and environment, which was developed for use in LMICs [32].
- Alzheimer’s Disease Cooperative Study-Activities of Daily Living Scale, a 23-item scale that addresses basic and instrumental activities of daily living and has been culturally adapted for use in Brazil [33,34]. This scale was used in Brazil only.
- Scale for the Instrumental Activities of Daily Living in the Elderly, an 11-item scale developed in South India that addresses cognitive and physical disability [35]. This scale was used in India only.

Family caregivers completed the following:

- The Zarit Burden Interview (ZBI), a widely used 22-item self-report measure of strain and stress [36].
- Dementia Caregiver Experience Scale, a 17-item measure to assess stress and strain (personal communication by Vaitheswaran, 2023), was included due to previous issues with the cross-cultural validity of the ZBI, which has been found to underestimate burden in LMIC settings [37].

Semistructured Interviews
To gather in-depth information about intervention acceptability, feasibility, and experiences of implementation, semistructured interviews were conducted with stakeholders. People with dementia and caregivers from the first 2 vCST groups in Brazil (12 dyads) and all vCST groups in India (15 dyads) were invited to take part in dyadic interviews. A convenience sample of group leaders from Brazil and India (n=7) and organizational decision makers (n=2, India only) was invited to take part in one-to-one interviews.

Interviews with people with dementia and caregivers were conducted in the language of vCST group instruction (English,
Tamil, or Brazilian Portuguese) and were led by a researcher from the respective institution who had not facilitated the group to reduce response bias. Interviews with group leaders and organizational decision makers were conducted in English. To reduce response bias, these were led by UK-based researchers who were not members of the immediate project team. All interviews took place over videoconferencing software.

On the basis of constructs from the CFIR, the interview guide was developed by researchers and clinical psychologists, with questions relating to the participants’ experience of taking part in the vCST group, experience of using an online platform, and barriers and facilitators to involvement.

Analysis

Quantitative

Primary analysis was descriptive and assessed recruitment, retention, and outcome measure completion. The pre- and postintervention outcome measures, that is, means and mean differences (calculated through paired 2-tailed *t* tests), were reported descriptively. Effect sizes were calculated using Cohen *d*. Statistical analysis was performed using SPSS (version 29; IBM Corp). All data were pseudoanonymized with a unique participant identification number.

Qualitative

Transcripts in Brazilian Portuguese or Tamil were translated before analysis. First, the researchers read the transcripts for data familiarization. We used a framework analysis approach to code the transcripts using inductive thematic analysis [38] and mapped inductive themes onto the CFIR. This enabled us to capture themes that were not covered by the CFIR. Transcripts were coded independently by 2 researchers using NVivo software (Lumivero), who met regularly to ensure that they were approaching the data in a similar way and then agreed on a final coding framework. Any discrepancies were examined and resolved through discussion. An exception is the interview transcripts of caregivers in India, as it was coded by an individual reviewer using Atlas software (Observational Health Data Sciences and Informatics) and discussed with a second researcher.

Ethical Considerations

Ethics approval was granted by the relevant body in each country. In Brazil, an ethics amendment was granted by the Federal University of Rio de Janeiro Institute of Psychiatry research ethics committee (ref: 57019616.5.1001.5263). In India, approval was granted by the institutional ethics committee at Schizophrenia Research Foundation; SCARF: Chennai (SRF-CR/17/0CT-2020). Informed consent was received from people with dementia and their caregivers upon recruitment. Participants did not receive compensation.

Interviews were recorded with consent and transcribed. All transcripts were pseudoanonymized with a unique participant ID number, and any identifying information was removed during transcription.

Results

Feasibility Outcomes

Recruitment Rate

A total of 59 participants were recruited to vCST groups, which was 91% of the target figure of 65 participants. Participant demographics are outlined in Table 1. In Brazil, 12 group sessions took place with a total of 44 participants between April 2021 and November 2022. In India, 5 group sessions were conducted with 15 participants between April 2021 and February 2022. This represents an average of 3.6 participants per group across both sites.
Table 1. Participant demographics.

<table>
<thead>
<tr>
<th></th>
<th>Brazil (n=44)</th>
<th>India (n=15)</th>
<th>Total (N=59)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Person with dementia</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Region, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southeast Brazil</td>
<td>41 (93)</td>
<td>—</td>
<td>41 (69)</td>
</tr>
<tr>
<td>South Brazil</td>
<td>2 (5)</td>
<td>—</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Northeast Brazil</td>
<td>1 (2)</td>
<td>—</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Chennai (India)</td>
<td>—</td>
<td>15 (100)</td>
<td>15 (25)</td>
</tr>
<tr>
<td><strong>Age (y), mean (range)</strong></td>
<td>77.1 (61-93)</td>
<td>77.3 (65-93)</td>
<td>77.2 (61-93)</td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>10 (23)</td>
<td>10 (67)</td>
<td>20 (34)</td>
</tr>
<tr>
<td>Female</td>
<td>34 (77)</td>
<td>5 (33)</td>
<td>39 (66)</td>
</tr>
<tr>
<td><strong>Ethnicity, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Asian</td>
<td>—</td>
<td>15 (100)</td>
<td>15 (25)</td>
</tr>
<tr>
<td>Black</td>
<td>3 (7)</td>
<td>—</td>
<td>3 (5)</td>
</tr>
<tr>
<td>Mixed</td>
<td>6 (14)</td>
<td>—</td>
<td>6 (10)</td>
</tr>
<tr>
<td>White</td>
<td>35 (80)</td>
<td>—</td>
<td>35 (59)</td>
</tr>
<tr>
<td><strong>Education (y), mean (range)</strong></td>
<td>11.5 (4-20)</td>
<td>13.0 (10-17)</td>
<td>11.9 (4-20)</td>
</tr>
<tr>
<td><strong>Type of dementia, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alzheimer disease</td>
<td>22 (50)</td>
<td>11 (73)</td>
<td>33 (60)</td>
</tr>
<tr>
<td>Vascular dementia</td>
<td>4 (9)</td>
<td>3 (2.0)</td>
<td>7 (12)</td>
</tr>
<tr>
<td>Mixed dementia (Alzheimer disease and vascular dementia)</td>
<td>1 (2)</td>
<td>1 (7)</td>
<td>2 (3)</td>
</tr>
<tr>
<td>Parkinson-related dementia</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Dementia with Lewy bodies</td>
<td>1 (2)</td>
<td>0 (0)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Variant unknown</td>
<td>16 (36)</td>
<td>0 (0)</td>
<td>16 (27)</td>
</tr>
<tr>
<td><strong>Caregiver</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sex, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>7 (16)</td>
<td>3 (20)</td>
<td>10 (17)</td>
</tr>
<tr>
<td>Female</td>
<td>37 (84)</td>
<td>12 (80.0)</td>
<td>49 (83)</td>
</tr>
<tr>
<td><strong>Age (y), mean (range)</strong></td>
<td>52.5 (32-71)</td>
<td>53.2 (29-72)</td>
<td>52.7 (29-72)</td>
</tr>
<tr>
<td><strong>Relationship to person with dementia, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spouse</td>
<td>8 (19)</td>
<td>4 (27)</td>
<td>12 (20)</td>
</tr>
<tr>
<td>Daughter or son</td>
<td>32 (74)</td>
<td>11 (73)</td>
<td>43 (73)</td>
</tr>
<tr>
<td>Daughter-in-law or son-in-law</td>
<td>1 (2)</td>
<td>0 (0)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>Other relative</td>
<td>2 (5)</td>
<td>0 (0)</td>
<td>2 (3)</td>
</tr>
<tr>
<td><strong>Living with person with dementia, n (%)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>11 (30)</td>
<td>2 (13)</td>
<td>13 (22)</td>
</tr>
<tr>
<td>Yes</td>
<td>26 (70)</td>
<td>13 (87)</td>
<td>39 (66)</td>
</tr>
</tbody>
</table>

\[^a\] Not applicable.
\[^b\] Missing data for Brazil: caregiver age, n=7.
\[^c\] Missing data for Brazil: relationship to person with dementia, n=1.
\[^d\] Missing data for Brazil: living with person with dementia, n=7.
Attendance
In Brazil, 52% (23/44) of participants attended all 14 sessions, and in India, 53% (8/15) of participants had full attendance.

Attrition
In Brazil, the attrition rate was 9% (4/44), denoting the percentage of participants who did not complete the vCST program due to various reasons: caregiver unavailability to support the participant (n=2, 50%), hospitalization due to COVID-19 (n=1, 25%), and to go on a vacation (n=1, 25%). There were no dropouts from vCST groups in India.

Outcome Measure Completion
Researchers completed preintervention outcome assessments with all people with dementia; however, some caregivers (3/59, 5%) were unavailable to provide preassessment measures. Retention of people with dementia to follow-up was 89% (39/44) in Brazil and 93% (14/15) in India. In India, 87% (13/15) of caregivers completed all follow-up assessments, but this figure was lower in Brazil (31/44, 70%). This was attributed to the caregivers being occupied with family and work commitments, particularly at a time of increased pressure during the pandemic. In addition, some people with dementia did not have 1 named caregiver and were supported by many family members or paid caregivers who did not always feel that they could provide accurate information. Overall, 68% (40/59) of participant dyads across both sites completed all postintervention outcome measures. No measures caused distress, and no measures had individual items missing.

Pre- and postintervention means and mean differences are outlined in Table 2. Analyses were exploratory and not powered to detect specific changes. The results suggest a small decrease in cognitive ability from baseline to follow-up. Small improvements across quality-of-life domains were observed in people with dementia. We observed moderate reductions in the activity of daily living ability across all domains in both sites. Conflicting outcomes were observed in caregiver burden outcomes, with a small reduction in burden scores on the ZBI but an increase in burden scores according to the Dementia Caregiver Experience Scale measure.

Multimedia Appendix 1 presents results by country. The direction and magnitude of change were similar across both sites; however, notable differences emerged: cognition where the decrease was smaller in India, quality-of-life score (social relationships) where the increase was smaller in India, and quality-of-life score (psychological and environment domains) where a reduction was observed in India compared with an improvement in Brazil.
### Qualitative Results

#### Overview

A total of 36 qualitative interviews were conducted. In Brazil, 12 people with dementia and their caregivers took part. In India, interviews were conducted with 15 people with dementia and caregivers. This comprises all participants from the first 2 vCST groups in Brazil, and all participants from the 5 groups in India. In addition, 4 group leaders from Brazil took part in interviews, and from India, 3 group leaders and 2 organizational decision makers from the NGO in India. All participants who were invited to the interviews agreed to take part.

Guided by the CFIR, we explored 2 main areas in the analysis: (1) acceptability of vCST and (2) barriers and facilitators to implementation.

#### Acceptability of vCST

All interview participants were asked directly about their experiences of taking part in vCST and were asked to reflect on how it compared to previous face-to-face activities. Overall, participant and caregiver evaluation of vCST was positive, with key benefits relating to providing occupation, enjoyment, and social interaction at the time of isolation:

> I liked her activeness and purposefulness...that itself is important. Earlier she used to simply sit but now she has something to do, so that kind of purposefulness is really appreciable. [Caregiver 4, India]

> We talk and such in the house, but we are only a few here. Now my family is almost just me and him...[the] television doesn’t interact. [Caregiver 8, Brazil]

### Table 2. Pre- and postintervention outcome measures.a,b

<table>
<thead>
<tr>
<th>Outcome (range)</th>
<th>Preintervention measures</th>
<th>Postintervention measures</th>
<th>Mean improvement (pretest-posttest)</th>
<th>Effect size (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Values</td>
<td>n (%)</td>
<td>Values, mean (SD)</td>
<td>Values</td>
<td>n (%)</td>
</tr>
<tr>
<td>ADAS-Cogc (0-70)</td>
<td>59 (100)</td>
<td>27.11 (12.92)</td>
<td>52 (88)</td>
<td>27.36 (14.53)</td>
<td>52 (88)</td>
</tr>
<tr>
<td>WHOQOL-BREFb: physical health (4-20)</td>
<td>59 (100)</td>
<td>14.48 (2.82)</td>
<td>51 (86)</td>
<td>14.80 (2.56)</td>
<td>51 (86)</td>
</tr>
<tr>
<td>WHOQOL-BREF: psychological (4-20)</td>
<td>59 (100)</td>
<td>14.17 (2.13)</td>
<td>51 (86)</td>
<td>14.75 (2.02)</td>
<td>51 (86)</td>
</tr>
<tr>
<td>WHOQOL-BREF: social relationships (4-20)</td>
<td>59 (100)</td>
<td>15.12 (1.71)</td>
<td>51 (86)</td>
<td>15.48 (1.96)</td>
<td>51 (86)</td>
</tr>
<tr>
<td>WHOQOL-BREF: environment (4-20)</td>
<td>59 (100)</td>
<td>15.03 (1.96)</td>
<td>51</td>
<td>15.25 (1.82)</td>
<td>51 (86)</td>
</tr>
<tr>
<td>ADCS-ADL.e (0-78)</td>
<td>44 (100)</td>
<td>44.34 (16.55)</td>
<td>40 (91)</td>
<td>42.00 (16.44)</td>
<td>40 (91)</td>
</tr>
<tr>
<td>IADL-EDRf—cognitive domain (0-100)</td>
<td>15 (100)</td>
<td>37.65 (19.77)</td>
<td>14 (93)</td>
<td>43.97 (20.72)</td>
<td>14 (93)</td>
</tr>
<tr>
<td>IADL-EDR—physical domain (0-100)</td>
<td>15 (100)</td>
<td>4.20 (7.95)</td>
<td>14 (93)</td>
<td>10.57 (15.71)</td>
<td>14 (93)</td>
</tr>
<tr>
<td>ZBIg (0-88)</td>
<td>56 (95)</td>
<td>35.02 (18.04)</td>
<td>47 (80)</td>
<td>32.91 (17.69)</td>
<td>46 (78)</td>
</tr>
<tr>
<td>DemCarEs.h (17-51)</td>
<td>53 (90)</td>
<td>28.94 (6.11)</td>
<td>46 (78)</td>
<td>29.35 (7.28)</td>
<td>43 (73)</td>
</tr>
</tbody>
</table>

---

aPositive maximum scale scores: ADAS-Cog=0, WHOQOL-BREF (Physical health, Psychological, Social relationships, Environment)=20, ADCS-ADL=78, IADL-EDR (Cognitive domain, Physical domain)=0, ZBI=0, DemCarES=17.

bEffect size was calculated using Cohen d (complete case analysis). No adjustments were made for multiple testing because analyses are exploratory.

cADAS-Cog: Alzheimer’s Disease Assessment Scale-Cognitive Subscale.

dWHOQOL-BREF: World Health Organization Quality of Life Brief Version.

eADCS-ADL: Alzheimer’s Disease Cooperative Study-Activities of Daily Living Scale, administered in Brazil only.

fIADL-EDR: Instrumental Activities of Daily Living for elderly people, administered in India only.

gZBI: Zarit Burden Interview.
hDemCarES: Dementia Caregiver Experience Scale.
At first I didn’t want to attend the sessions (laughs). I fought, I wanted to hit everyone, but I liked it. [Person with dementia 1, Brazil]

The remote delivery and national recruitment in Brazil also enabled the attendance of some participants from outside the urban centers of Rio de Janeiro and São Paulo, where most services are provided:

You’re interacting there from Rio, [name of another participant] there from Itapetininga, the other lady also from another place...with this pandemic business...we don’t need to have physical contact. I think it’s great. [Caregiver 8, Brazil]

However, many, in particular, the facilitators who had had the experience of delivering both vCST and face-to-face CST, felt that the social connection and stimulation would have been stronger if the intervention had taken place face to face:

There are many more activities that can be done in person, rather than virtually...like for example, throwing ball to each other, doing physical activities together. Even sensory stimulation like...hearing sounds or seeing things...And I feel just physically being present and seeing other people is definitely...much more helpful. [Group leader 3, India]

I think it would have been better if it could have happened in person. But during the COVID situation...this was more helpful and comfortable as anybody can attend from any place. Maybe still, I feel it would have been more beneficial for the dementia group if it were a direct session. [Caregiver 6, India]

The participants observed additional issues with intervention acceptability that were related to the participants’ access to suitable technology and computer literacy, which was compounded by cognitive impairment:

The main issue was internet. I would say...so we had only three participants in a group...along with a facilitator and a co-facilitator...which means that, like five different internet connections. So, the problem was if even one participant had a disruption in their internet, it tends to affect the whole group. [Group leader 1, India]

At first it was more difficult, because the computer she could use at this time, I was using for work...so she had to do it on her phone...The images were too small for her to see, so that got in the way. [Caregiver 7, Brazil]

Group leaders also reflected that it was more difficult to gauge engagement and facilitate a group virtually, as opposed to face to face:

Just knowing the body language, if the person is feeling sleepy, or the person’s not enjoying it and stuff like that. You’re not able to notice it as much because it is a virtual setup. [Group leader 1, India]

Sometimes...the participants would talk over other people. We will ask someone a question, and that person...would answer, but then another person would answer also, and the two answers were colliding there, and it was hard to manage that, because it was virtual sessions. [Group leader 2, Brazil]

Facilitators and Barriers to Implementation

Facilitators and barriers are included in Tables 3 and 4, categorized by CFIR domain with illustrative quotes. Key facilitators included the following:

- **Innovation**: Facilitators included the evidence base of CST and its advantage over other psychosocial interventions as a manualized intervention, which was also flexible to the needs of the participants. Some group leaders reflected that they were aware of few other interventions for people with dementia taking place virtually at the time.

- **Outer setting**: An international collaborative effort enabled funding and sharing of protocols and training materials. Many caregivers reflected that they were appreciative that the person with dementia could attend vCST at a time of social isolation due to COVID-19 restrictions.

- **Inner setting**: Staff in both sites were motivated to offer a service for people with dementia, and many participants reflected on the need for more support for people with dementia. Another facilitator to implementation was the training and supervision of staff at the NGO and trainee psychologists at the university. The NGO in India were able to appoint permanent staff members to take on vCST responsibilities as part of their role and integrate vCST into the existing services and caseload.

- **Individuals**: Most people with dementia relied on caregivers’ support and would often miss sessions if their caregiver was unavailable. All groups also required 2 group leaders: one to lead the activities and another to provide technological support and to contact caregivers if a participant was struggling to engage. In India, group leaders reflected that adoption of vCST improved if it was suggested to participants by a clinician.

- **Processes**: Key implementation strategies included providing mock vCST sessions with caregivers and people with dementia to orient them to the platform and posting out activity packs to those who did not have resources at home.
<table>
<thead>
<tr>
<th>CFIR® domain and subdomains&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Innovation</strong></td>
<td></td>
</tr>
<tr>
<td>Innovation evidence-base</td>
<td>“In terms of evidence based published literature information… the effectiveness of CST and the cost effectiveness in other centers… That helped in choosing the most appropriate intervention.” [Decision maker 1, India]</td>
</tr>
<tr>
<td>Innovation relative advantage</td>
<td>“There was this one organization… a day center facility were doing… one-on-one video calls to have some sort of a social interaction during the pandemic.” [Group leader 2, Brazil]</td>
</tr>
<tr>
<td>Adaptability (of vCST&lt;sup&gt;c&lt;/sup&gt; protocol)</td>
<td>“I think we had flexibility, because as I said one was the education level of patients and then the language that had to be used.” [Decision maker 2, India]</td>
</tr>
<tr>
<td><strong>Outer setting</strong></td>
<td></td>
</tr>
<tr>
<td>Local conditions (need for socializing during lockdowns)</td>
<td>“He was… looking forward to the session, especially social interaction because the pandemic had obviously you know sort of cut down a lot of such interactions.” [Caregiver 1, India]</td>
</tr>
<tr>
<td>Partnerships and connections (international research partnership)</td>
<td>“We based it ourselves in this protocol, which was already published with some guidelines for developing the CST virtually.” [Group leader 1, Brazil]</td>
</tr>
<tr>
<td>Financing (international research funding)</td>
<td>“We were able to purchase the items that we need to deliver CST at our center. And for regarding technology… we were able to provide some of the participants with a tablet computer and the data for them.” [Decision maker 1, India]</td>
</tr>
<tr>
<td><strong>Inner setting</strong></td>
<td></td>
</tr>
<tr>
<td>Tension for change (need for psychosocial treatment)</td>
<td>“There is no actual evidence based structured manual intervention available in India prior to this, so this provided as an opportunity to make it available for our patients.” [Decision maker 1, India]</td>
</tr>
<tr>
<td>Culture—learning centeredness (supporting trainee psychologists, Brazil)</td>
<td>“I really like participating on the project from my experience, in gaining experience, on like clinical experience and also a little bit of research too.” [Group leader 3, Brazil]</td>
</tr>
<tr>
<td>Compatibility (with service and caseload, India)</td>
<td>“We have a regular clinic so we identify participants from the clinic.” [Decision maker 1, India]</td>
</tr>
<tr>
<td>Access to knowledge and information (training and supervision)</td>
<td>“We had training, of course, and we also had regular supervision from our supervisor.” [Group leader 1, India]</td>
</tr>
<tr>
<td>Work infrastructure—staff (at NGO&lt;sup&gt;d&lt;/sup&gt;, India)</td>
<td>“Making sure that the facilitators are in substantive posts and not in fleeting positions so they are available for a longer time.” [Decision maker 1, India]</td>
</tr>
<tr>
<td><strong>Individuals</strong></td>
<td></td>
</tr>
<tr>
<td>Opinion leaders (recommendation from doctors)</td>
<td>“If the doctor sometimes says, ‘you should do this, this will be beneficial for you,’ it really helps in the Indian context of the doctor’s word for you.” [Group leader 2, India]</td>
</tr>
<tr>
<td>Other implementation support—availability or capability (caregivers)</td>
<td>“Some of [the caregivers] would stay next to the person living with dementia… especially when the person was a little bit shy, [or] had more difficulty with technology… They were… mediating this communication.” [Group leader 2, Brazil]</td>
</tr>
<tr>
<td>Other implementation support—availability or capability (cofacilitator)</td>
<td>“One of the psychologists is delivering the session, and we need someone to support us at the technical end, we need someone to support us.” [Group leader 2, India]</td>
</tr>
<tr>
<td>Intervention recipient—need (person with dementia—need to stay home and subsequent isolation)</td>
<td>“Some of these people would not have come for in-person CST, because they could not afford transportation, did not have proper transportation, were frail, or had some kind of physical comorbidity or pain.” [Group leader 2, India]</td>
</tr>
<tr>
<td>CFIR(^a) domain and subdomains(^b)</td>
<td>Quotes</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td><strong>Implementation process</strong></td>
<td></td>
</tr>
</tbody>
</table>
| Tailoring strategies (mock vCST sessions and activity packs) | • “We do have one trial session, where I sit with them individually. And then we have one group trial session, to see if they’re comfortable in a group.” [Group leader 2, India]  
• “We posted the materials...for number games, we had paper sheets. And colouring papers and some origami papers...We took printouts and posted it to their house.” [Group leader 1, India] |

\(^a\)CFIR: Consolidated Framework for Implementation Research.  
\(^b\)Context-specific descriptions are given in parentheses.  
\(^c\)vCST: virtual cognitive stimulation therapy.  
\(^d\)NGO: nongovernmental organization.
<table>
<thead>
<tr>
<th>CFIR(^a) domain and subdomains(^b)</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Innovation</strong></td>
<td></td>
</tr>
<tr>
<td>Adaptability (virtual delivery of CST(^c))</td>
<td>“I was running face-to-face sessions before they started [vCST]. Face-to-face CST was great...my group ran with eight members, six to eight, consistently. So, I had a huge group coming every Friday. It was amazing, they could form more connections, and turn taking is a little bit easier...It’s a little harder like you with the Zoom.” [Group leader 2, India]</td>
</tr>
<tr>
<td>Innovation design (need for marketing)</td>
<td>“It doesn’t have much publicity. If it wasn’t for chance, if this person hadn’t put us in, I wouldn’t have made it. So, I think in terms of dissemination it could be broader.” [Caregiver 4, Brazil]</td>
</tr>
<tr>
<td><strong>Outer setting</strong></td>
<td></td>
</tr>
<tr>
<td>Critical incidents (COVID-19 pandemic)</td>
<td>“It was COVID and people are falling sick...even the facilitators are sick, at some point.” [Group leader 1, India]</td>
</tr>
<tr>
<td>Local attitudes (dementia awareness)</td>
<td>“In Brazil, I think it’s a cultural thing to think that dementia symptoms it’s part of a natural aging...So, when older people, and people living with dementia...come to a doctor to be evaluated they sometimes don’t have mild symptoms anymore.” [Group leader 2, Brazil]</td>
</tr>
<tr>
<td>Local attitudes (traditional focus on medical model)</td>
<td>“People weren’t aware of psychosocial interventions for dementia prior to this. They had very different model for working with people with dementia.” [Decision maker 1, India]</td>
</tr>
<tr>
<td>Local conditions (access to technology)</td>
<td>“Most of the people that we had in the groups were from the south eastern region. And that’s kind of a more developed region financially...I think today most people in Brazil have access to internet. Maybe not their computer, but maybe cell phones and something like that.” [Group leader 3, Brazil]</td>
</tr>
<tr>
<td><strong>Inner setting</strong></td>
<td></td>
</tr>
<tr>
<td>Structural characteristics—work infrastructure (staff availability)</td>
<td>“When we think of scaling it up, we might have to do it first of all in institutions where there is enough manpower of mental health professionals to deliver the CST...dementia care in India is still mental health care and we’re still very under-resourced as far as manpower is concerned.” [Decision maker 2, India]</td>
</tr>
<tr>
<td><strong>Individuals</strong></td>
<td></td>
</tr>
<tr>
<td>High-level leaders—capability (lack of dementia awareness)</td>
<td>“Some of the policymakers, who we interviewed at the beginning [in previous stakeholder engagement] weren’t even aware of the issues relating to dementia.” [Decision maker 1, India]</td>
</tr>
<tr>
<td>Intervention deliverers—capability (basics in clinical skills needed)</td>
<td>“I think we if we didn’t have the training, it would be very hard to just come to the groups...I didn’t have any contact [with people with dementia] before.” [Group leader 3, Brazil]</td>
</tr>
<tr>
<td>Intervention recipients—capability (sensory impairment and computer literacy)</td>
<td>“So, one challenge was delivering it virtually. My mother was not able to hear very well. Now she has a hearing aid, she has the headphones but still that was a part of a problem of communication.” [Caregiver 5, India]</td>
</tr>
<tr>
<td></td>
<td>“I don’t know how to use the computer (laughs).” [Person with dementia 10, Brazil]</td>
</tr>
<tr>
<td><strong>Implementation process</strong></td>
<td></td>
</tr>
<tr>
<td>Assessing needs—innovation recipients (severity of dementia)</td>
<td>“If you have some difference in severity of dementia, because the activities demand something, and maybe it can be boring for who is not so severe.” [Group leader 4, Brazil]</td>
</tr>
<tr>
<td>Assessing needs—innovation recipients (baseline assessments)</td>
<td>“The first is, I think, the baseline evaluations were very long, and that was kind of hard on the, not on the people with dementia, but on their family members, the caregivers.” [Group leader 3, Brazil]</td>
</tr>
<tr>
<td>Reflecting and evaluating—implementation (lack of long-term follow-up)</td>
<td>“One question that most people with dementia their caregivers made was, if it was possible to have more than 14 sessions. So maybe adapting the maintenance CST for the virtual program. I think it would be a suggestion for the future.” [Group leader 1, Brazil]</td>
</tr>
</tbody>
</table>

\(^a\)CFIR: Consolidated Framework for Implementation Research.  
\(^b\)Context-specific descriptions are given in parentheses.  
\(^c\)CST: cognitive stimulation therapy.
Key barriers related to the 5 CFIR domains are as follows:

- **Innovation.** Most group leaders highlighted challenges with the online delivery of CST in terms of facilitating a group effectively, meeting individual needs, and supporting participants with the videoconferencing platform. Many leaders reflected on the comparative ease of facilitating a group in person. These issues are outlined in detail in the Acceptability of vCST section. Finally, group leaders and caregivers highlighted the need for marketing to raise awareness of vCST.

- **Outer setting.** While COVID-19 necessitated and possibly facilitated the online delivery of CST, staff and participant illness during the pandemic was a barrier to group delivery and attendance. Staff at both sites reflected on a lack of dementia awareness, resulting in participants presenting later to clinical services, which is a barrier to recruiting participants with mild to moderate dementia. Similarly, group leaders and decision makers reflected on a lack of awareness of psychosocial interventions for dementia, with the medical model tending to prevail. Finally, in both sites, it was highlighted that poor or limited access to technology is a barrier to involvement.

- **Inner setting.** The limited availability of mental health personnel was highlighted as a barrier to the wider scale-up of vCST in India.

- **Individuals.** People with dementia and caregivers faced barriers to taking part in vCST, including a lack of computer literacy and sensory impairment that impacted engagement. Organizational decision makers in India reflected on the lack of dementia awareness within high-level policy makers.

- **Processes.** One group leader reflected on the length and burden of the baseline assessments on people with dementia and caregivers. Many people with dementia and caregivers expressed a wish for the vCST groups to continue beyond the 14 sessions. At the NGO in India, it was possible to follow up with patients on the caseload; however, group leaders in Brazil wished to be able to continue to support participants and caregivers.

**Discussion**

**Principal Findings**

We found that it was feasible and acceptable to deliver CST virtually in Brazil and India. We recruited 91% (59/65) of the target sample and were able to run 17 vCST groups. Attrition was low (4/49, 7%), and attendance was moderate, with 53% (31/59) of participants attending all 14 sessions. This is in contrast to a previous trial of face-to-face CST in Brazil, where attrition was similar (6%) but attendance was high (mean 12.8 sessions, median 14 sessions) [39]. In a previous pilot study of face-to-face CST in Chennai, India, attrition was higher with 3 out of 9 participants dropping out [16]. However, these comparisons should be interpreted with care due to small sample sizes and the impact of COVID-19 in both countries.

Outcome measure completion was slightly lower than the target of 75%, as only 68% (40/59) of participant dyads completed all follow-up outcome measures, suggesting a possible measurement burden. Small improvements across all quality-of-life domains were observed in people with dementia. All results should be interpreted with care, as the study was not controlled. Any changes cannot be ascribed to the vCST intervention specifically, and the impact of COVID-19 and consequent social isolation may have played a role in pre- and postintervention measurement changes.

There were some differences in the barriers and facilitators to implementation across the 2 sites. vCST was delivered in an NGO in Chennai, where participants could be recruited from the patient caseload. In Brazil, vCST was delivered through a university where recruitment took place from the community and memory clinics and NGOs who were partnered with the study. Although staff turnover was a barrier to implementation in the NGO in India, decision makers reflected that it would be possible to build CST or vCST into services due to its compatibility with current ways of working. This may have been more of a challenge in the university setting in Brazil, where there is no patient caseload or clinic infrastructure.

However, there is scope to build partnerships with community organizations and clinics to recruit participants. Upskilling trainee psychologists to deliver vCST in Brazil also presents a low-cost and scalable solution to implementing vCST in a university setting. Similar solutions have been used for other psychosocial interventions for dementia in other countries [40,41].

A major barrier in both sites was poor or limited access to technology and computer literacy. This issue was also highlighted in studies of vCST conducted in the United Kingdom [23,24]. A survey of the digital divide in India found that just 38% of households are digitally literate [42]. Access is higher in Brazil, where 80% of households have internet access [43]. In both countries, digital access intersects with age, gender, education level, and ethnicity [42-44]. While virtual interventions provide service access to those living in geographically isolated locations, the digital divide is greater in rural areas; in Brazil, only 53% of the rural population have internet access, while 88% in urban areas have internet access [43], and the figures are lower overall in India where in rural areas, only 31% of the population use the internet, while in the urban areas, the percentage rises to 67% [42]. To overcome the barriers to technology access in India, the NGO loaned tablets to participants, which required sufficient funding and resources. We addressed the issues related to digital literacy by implementing the following measures: (1) group leaders provided mock vCST sessions to familiarize participants with the videoconferencing platform, (2) a coleader was available specifically for technology support, and (3) groups sizes were smaller so that all participants could be viewed on the screen at once (average 3.6 participants compared with 6-8 according to the original CST protocol [12]).

Most people with dementia were reliant on caregivers’ technical support to use the videoconferencing technology, and in some cases, caregivers remained present throughout the group sessions. This raises a key issue for those without caregivers, who could be systematically excluded from taking part in virtual psychosocial interventions. If vCST were implemented as a dyadic intervention, this could improve caregivers’ awareness of dementia and person-centered approaches, which is important...
given the limited number of dementia awareness programs in LMIC settings [45]. However, it could also negatively impact the engagement of the person with dementia, as one of the proposed mechanisms of action of CST is the supportive learning environment, where people with dementia support each other without judgment or embarrassment [46]. If vCST is delivered dyadically, we recommend that participants are briefed at the start of the program to set expectations about the caregivers’ level of involvement in the vCST sessions and that people with dementia are provided with opportunities to take part in activities and discussions alone. Further research could explore the impact of dyadic delivery on outcomes for people with dementia and their caregivers.

Limitations

In both sites, it is likely that the sample was not representative of the broader population of people living with dementia and their caregivers. Specifically, in Brazil, the sample comprised mostly White individuals (35/44, 80%), which does not reflect the majority Black and mixed Brazilian population. Most participants were from the urban region of southeast Brazil, although the remote method of recruitment did enable participation from areas outside of this region, which were underserved in terms of research and clinical services. In India, all participants were recruited from the same region and were already attending clinical services; this might have resulted in a sample skewed to those with the means to access services.

Online delivery may result in a self-selecting sample, who are more likely to be educated to a higher level and more affluent than the broader population. The mean number of years of schooling of our sample was 11.5 (SD 1.2) in Brazil and 13.0 (SD 2.6) in India. This compares to a national average of 2.5 in Brazil and 1.4 in India for the population aged ≥25 years in 1970 and 1971, respectively [47] although there is huge regional variation in education levels in both countries. To overcome issues related to digital exclusion in India, tablets were loaned to those who needed them. However, in Brazil, people without access to their own technology were excluded.

In terms of the qualitative component, most themes and quotes from a participant perspective were collected from caregivers rather than people with dementia. This is because cognitive impairment affected their recall of sessions. Despite this, caregivers and group leaders reflected on the perceived participant experience of vCST sessions. Interviews took place with all caregiver dyads from India, but only those from the first 2 vCST groups in Brazil due to staff availability. However, the reflections from group leaders and organizational decision makers relate to all vCST groups. Interviews with group leaders and decision makers were carried out in English by a UK-based researcher who was not a member of the immediate research team. This was to limit response bias and encourage honest and critical feedback; however, it limited the interview to people who speak English and may have compromised the representation of non-English speakers.

Finally, the vCST intervention was tested in 2 sites, a university in Brazil and an NGO in India, resulting in lessons for implementation that could be explored in other sites and countries; nevertheless, we acknowledge the limited generalizability of these findings.

Future Research

To date, vCST has only been trialed within a pandemic context. While this was acceptable to participants during a time of social isolation, many caregiver dyads and group leaders expressed a desire for CST to take place in person. Future research could explore the feasibility of vCST outside of the pandemic context, perhaps specifically targeting those who cannot access in-person services due to limited mobility, health issues, or geographic isolation.

In addition, although there is a strong evidence base for in-person CST, we do not know if the benefits to cognition and quality of life are conferred to the same level over online delivery. A recent feasibility study of vCST (in press Spector, 2023) has shown that a full-scale randomized controlled trial is warranted.

Conclusions

The 14-session vCST program for people with dementia was successfully trialed in a university setting in Brazil and in an NGO in Chennai, India. vCST offered a feasible alternative to in-person groups during the period of pandemic restrictions with potential benefits to quality of life, but there were barriers related to technology access and computer literacy. Outside of the pandemic context, vCST could be provided to people with dementia who are geographically isolated or who have mobility- or health-related difficulties.

Acknowledgments

The authors would like to thank all the participants of the virtual cognitive stimulation therapy (vCST) groups and the vCST group leaders.

This work is supported by the following Global Alliance for Chronic Diseases (GACD) funding agencies: The United Kingdom Medical Research Council (MRC: MR/S004009/1) and the Indian Council of Medical Research (ICMR: Indo-foreign/67/M/2018-NCD-I). No funding bodies were involved in the design, collection, analysis, interpretation, or writing of the research or manuscript. The views expressed are those of the authors and not necessarily those of GACD, the MRC, or ICMR.
Conflicts of Interest

AS delivers occasional cognitive stimulation therapy (CST) training courses on a consultancy basis. DCM provides CST training in Brazil, and SV delivers CST training in India.

Multimedia Appendix 1

Pre- and postintervention outcome measures by country.

References


40. Livingston G, Barber I, Rapaport P, Knapp M, Griffin M, King D, et al. Clinical effectiveness of a manual based coping strategy programme (START, STrAtegies for RelaTives) in promoting the mental health of carers of family members with...


Abbreviations
CFIR: Consolidated Framework for Implementation Research
CST: cognitive stimulation therapy
LMIC: low- and middle-income countries
NGO: nongovernmental organization
vCST: virtual cognitive stimulation therapy
ZBI: Zarit Burden Interview

©Emily Fisher, Shreenila Venkatesan, Pedro Benevides, Elodie Bertrand, Paula Schimidt Brun, Céline El Baou, Cleusa P Ferri, Jane Fossey, Maria Jelen, Jerson Laks, Lisa Liu, Daniel C Mograbi, Nirupama Natarajan, Renata Naylor, Despina Pantoulı, Vaishnavi Ramanujam, Thara Rangaswamy, Raquel L Santos de Carvalho, Charlotte Stoner, Sridhar Vaitheswaran, Aimee Spector. Originally published in JMIR Aging (https://aging.jmir.org), 11.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
A Personalized and Interactive Web-Based Advance Care Planning Intervention for Older Adults (Koda Health): Pilot Feasibility Study

R Lynae Roberts1, PhD; Katelin D Cherry1, MBE; Desh P Mohan1, MD; Tiffany Statler2, DMS, PA; Eric Kirkendall3, MD, MBI; Adam Moses3, MHA; Jennifer McCraw3, RN, MSN, MBA; Andrew E Brown III3, MPH; Tatiana Y Fofanova1, PhD; Jennifer Gabbard4, MD

Corresponding Author:
R Lynae Roberts, PhD

Abstract

Background: Advance care planning (ACP) is a process that involves patients expressing their personal goals, values, and future medical care preferences. Digital applications may help facilitate this process, though their use in older adults has not been adequately studied.

Objective: This pilot study aimed to evaluate the reach, adoption, and usability of Koda Health, a web-based patient-facing ACP platform, among older adults.

Methods: Older adults (aged 50 years and older) who had an active Epic MyChart account at an academic health care system in North Carolina were recruited to participate. A total of 2850 electronic invitations were sent through MyChart accounts with an embedded hyperlink to the Koda platform. Participants who agreed to participate were asked to complete pre- and posttest surveys before and after navigating through the Koda Health platform. Primary outcomes were reach, adoption, and System Usability Scale (SUS) scores. Exploratory outcomes included ACP knowledge and readiness.

Results: A total of 161 participants enrolled in the study and created an account on the platform (age: mean 63, SD 9.3 years), with 80% (129/161) of these participants going on to complete all steps of the intervention, thereby generating an advance directive. Participants reported minimal difficulty in using the Koda platform, with an overall SUS score of 76.2. Additionally, knowledge of ACP (eg, mean increase from 3.2 to 4.2 on 5-point scale; \( P < .001 \)) and readiness (eg, mean increase from 2.6 to 3.2 on readiness to discuss ACP with health care provider; \( P < .001 \)) significantly increased from before to after the intervention.

Conclusions: This study demonstrated that the Koda Health platform is feasible, had above-average usability, and improved ACP documentation of preferences in older adults. Our findings indicate that web-based health tools like Koda may help older individuals learn about and feel more comfortable with ACP while potentially facilitating greater engagement in care planning.

(JMIR Aging 2024;7:e54128) doi:10.2196/54128

KEYWORDS

advance care planning; ACP; digital health tools; system usability; gerontology; geriatric; geriatrics; older adult; older adults; elder; elderly; older person; older people; ageing; aging; adoption; acceptance; usability; digital health; platform; website; websites

Introduction

Advance care planning (ACP) is a process by which individuals choose their goals of care, quality of life priorities, and potential future medical intervention preferences and then communicate these values [1-3]. Actions taken during ACP include choosing a surrogate decision-maker (SDM), completing advance directives, and discussing a patient’s wishes with loved ones and health care providers. With increasing average lifespans, ACP is a vital component of high-quality care to ensure that patients’ care when facing serious illness is concordant with their values and goals. Currently, anywhere between 3% to 47% of patients may receive medical care that is not consistent with the patient or their loved ones’ wishes [4-6]. Because of these inconsistencies, it’s estimated that US $75.7 billion to US $101.2 billion is spent on overtreatment or low-value care each year in the United States [7]. These findings highlight the need to increase communication regarding medical care planning.

ACP is associated with decreased anxiety among patients’ family and caregivers [8], improved patient quality of life [9],
decreased unwanted medical care [10-12], and decreased health care costs [13-16]. Despite the promising evidence of benefits [17], rates of ACP remain low, with many patients and families avoiding these discussions until the patient’s condition has deteriorated and is suboptimal for end-of-life decision-making. In addition, less than 11% of Medicare beneficiaries discuss ACP with their medical providers [18-21]. In the United States, approximately 37% or less of individuals report having some kind of advance directive, which could include a medical power of attorney or a living will [22,23]. ACP rates are often even lower within historically marginalized communities in America [24-28].

While the majority of patients express positive views or interest in ACP [22,29], they may not know how to begin the process. Initiation of ACP conversations seems to be a major barrier, as patients may be reluctant to broach the subject with their health care providers, while clinicians report having insufficient training or time to conduct ACP discussions [22,30,31] during busy clinic visits. Though patients believe that their medical providers should initiate conversations about ACP [32,33], 17% or fewer of patients report discussing ACP with their medical team [34]. Given the barriers for many patients in learning about and completing ACP, it is important to explore alternative approaches that may better support health care providers in facilitating this vital service [29]. Digital health tools could be a potential solution to improve equitable access to ACP for patients and to engage loved ones and health care providers in the process [35,36]. Scoping reviews have concluded that currently available web-based ACP programs are feasible and generally well-received by users [37,38], but the quality of the content greatly varies [39].

Current ACP online programs are primarily static web-based forms and do not include interactive educational content, plain language, or the capability to allow for official signing of ACP documentation [40]. Other ACP programs are geared toward specific patient populations and are therefore not generalizable to all individuals. Additionally, the sections on medical interventions in many online ACP resources fail to mention some common life support treatments that an individual may experience at end of life or with serious illness [41]. Several available smartphone apps also provide some education or actionable decision-making in regard to ACP, but currently available apps lack sufficient features or have poor functionality, limiting their practicality [42].

The Koda digital ACP platform seeks to fill the gaps found in current offerings by providing a highly interactive solution that is suitable for all technological knowledge levels and more inclusive of the most common life-support treatment options. The platform guides patients through a personalized, interactive guide, which includes video-based educational content and decision-making guides for patients and their loved ones. Users are able to select health care goals and indicate their wishes regarding potential future medical interventions, all of which auto-generate into easily accessible documentation that can be shared with loved ones and health care providers. Koda was created to help facilitate informed discussions of ACP and to provide a tool that patients could use freely and effectively without an added time burden for clinicians.

A previous retrospective report on Koda [43] was conducted with a sample of patients with serious illness. The findings of that quality improvement report showed that 53% of referred patients completed their ACP through the platform. However, due to the retrospective nature of this prior study, we were unable to assess self-report usability metrics or change in opinions or knowledge after platform use. The main objective of this study was to determine the reach, adoption, and usability of Koda to conduct digital ACP within a university health system.

**Methods**

**Population and Recruitment**

This pilot study included adult patients aged 50 years or older who had an active MyChart account, defined as one that was used within the past 90 days. Participants were excluded if they were younger than 50 years, were non–English speaking, had a diagnosis of Alzheimer disease or Alzheimer disease–related dementias, or had blindness based on electronic health record (EHR) International Classification of Diseases, 10th Edition codes.

An EHR algorithm was created to identify eligible participants (eg, filtering for the inclusion criteria) from the Atrium Health Wake Forest Baptist network (AHWFB). The EHR included data on age, gender, race, and diagnoses. The AHWFB is a large, quaternary health system affiliated with an accountable care organization program that incorporates more than 150 primary care and multispecialty practices with more than 330 physicians and advanced practice providers in 80 different locations in communities throughout central North Carolina.

Eligible participants were sent an electronic invitation to participate via Epic MyChart with an informational message about the study and an embedded hyperlink to the Koda platform. If a participant was interested, they were directed to a web-based consent form, and electronic informed consent was obtained. Participants were instructed to create a Koda Health account, complete a preassessment survey, complete using the Koda platform, and then complete a postassessment survey. Follow-up messages were sent 2 weeks later to any nonresponders. If a participant pressed “accept” but did not create a Koda account, they were approached 2 weeks later by the research team to facilitate the process.

**Ethical Considerations**

This study was approved by the Atrium Health Wake Forest Institutional Review Board (IRB00076779). All human subjects data were deidentified. Study participants were sent a US $25 gift card for completing the steps of the study.

**Intervention: Koda Digital Platform**

The Koda application was previously developed prior to this pilot study. Briefly, this occurred as an iterative process with input from ACP content experts (eg, geriatric and palliative medicine physicians), as well as input from end users surrounding their preferences, which was obtained by survey results. The patient-facing Koda ACP platform consists of 4 sections, focusing on values; individual definitions of quality...
of life; SDM preferences; and medical care preferences, with specific regard to cardiopulmonary resuscitation (CPR), mechanical ventilation, artificial nutrition via a feeding tube, and dialysis. This was in alignment with consensus recommendations on core components of ACP. Motivational interviewing techniques [44] were used to highlight the importance of planning and to motivate patients to communicate their wishes. Each section includes educational audio-video content and expandable information for additional questions users may have. For this study, platform completion progress was tracked through the Koda Health administration portal. Once the participant completed using the Koda platform by indicating their decisions within each section, their answers were autopopulated into a state-specific advance directive, which participants were able to sign or notarize online. Using the platform takes approximately 20-30 minutes. See Figure 1 for example displays directly from the platform.

**Figure 1.** Overview of the Koda platform interface.

### Outcome Measures

Primary and secondary outcome measures consisted of reach, adoption, and usability. Reach was defined as the proportion of eligible participants who clicked on the embedded hyperlink to the Koda application. Adoption was defined as the proportion of participants who completed using the Koda platform. Usability was measured using the validated System Usability Scale (SUS) [45,46], a 10-item questionnaire. Items are rated on a 5-point (ie, 0 to 4) scale and responses to all items are
summed and multiplied by 2.5. Possible scores range from 0 to 100, with scores of 68 or higher indicating above-average usability [47]. The SUS was reliable for the current sample with $\alpha = .87$, which is comparable to psychometric findings from multiple studies of the SUS (ranging from $\alpha = .83$ to $\alpha = .97$) [48].

Exploratory outcomes included ACP knowledge and readiness using an adapted, self-report ACP engagement survey [49] and were measured before and after using the Koda platform. To assess patients’ self-rated knowledge of ACP, 4 Likert-type questions were asked, which included “Do you know what advance care planning is?” “How well informed are you about who can be a medical decision maker?” “How well informed are you about what makes a good medical decision maker?” and “How well informed are you about the different amounts of flexibility a person can give their medical decision maker?” The answer options were on a scale from 1 to 5, with higher scores correlating to higher ACP knowledge. Reliability for the adapted ACP Engagement Survey was good ($\alpha = .82$) and comparable to previous psychometric research for other brief versions of this tool (ranging from $\alpha = .84$ to $\alpha = .97$) [49].

To assess patients’ readiness to engage with ACP, 4 Likert-type questions were asked, which included “What describes you best when it comes to your comfort level in thinking about your care if you become seriously ill?” “How ready are you to talk to your decision maker about the kind of medical care you would want if you were very sick or near the end of life?” and “How ready are you to talk to your healthcare provider about the kind of medical care you would want if you were very sick or near the end of life?” These answer options were on a scale from 1 to 5, with higher scores indicating greater readiness. The fourth question, “How worried are you when you think about your future illnesses?” was reverse scored so that higher scores would indicate less ACP readiness.

**Statistical Analyses**

Reach, adoption, and SUS scores were calculated and reported as percentages and percentiles, respectively. Success was defined a priori by having ≥40% of enrolled participants complete using the Koda platform and obtaining an above-average SUS score of 68 or higher [50-52]. For descriptive analyses, means and SDs were used to describe continuous variables and percentages and frequencies were used to describe categorical variables. Shapiro-Wilk, D’Agostino-Pearson, and Anderson-Darling tests were conducted to assess data normality, and study variables were found to have non-Gaussian distribution. Therefore, nonparametric methods were used for any inferential analyses. Changes from pre- to post-Koda use were analyzed with the 2-tailed Wilcoxon signed rank test with continuity correction. Analyses were performed using R (version 4.2; R Foundation for Statistical Computing). $P < .05$ was considered to be statistically significant.

**Results**

**Demographics**

Table 1 shows participant demographic and baseline characteristics. The mean age was 62.8 (SD 9.3) years, with ages ranging from 50-99 years. Of enrolled participants, 127 (78.9%) were female, and 46 (28.6%) were Black or African American. A total of 49 patients (30.4%) self-reported poor to fair health, 55 (35.2%) had a personal loss or misfortune in the past year, and 103 (66%) had been a caregiver of a loved one with a serious illness.
Table. Demographics of participants who created a Koda account.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Participants (n=161)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>62.8 (9.3)</td>
</tr>
<tr>
<td>Sex , n (%)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>127 (78.9)</td>
</tr>
<tr>
<td>Male</td>
<td>34 (21.1)</td>
</tr>
<tr>
<td>Ethnicity, n (%)</td>
<td></td>
</tr>
<tr>
<td>Hispanic or Latinx</td>
<td>1 (0.6)</td>
</tr>
<tr>
<td>Not Hispanic or Latinx</td>
<td>159 (98.8)</td>
</tr>
<tr>
<td>Not reported</td>
<td>1 (0.6)</td>
</tr>
<tr>
<td>Race, n (%)</td>
<td></td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>1 (0.6)</td>
</tr>
<tr>
<td>Black/African American</td>
<td>46 (28.6)</td>
</tr>
<tr>
<td>White</td>
<td>113 (70.2)</td>
</tr>
<tr>
<td>Self-reported health status, n (%)</td>
<td></td>
</tr>
<tr>
<td>Poor</td>
<td>10 (6.4)</td>
</tr>
<tr>
<td>Fair</td>
<td>39 (25)</td>
</tr>
<tr>
<td>Good</td>
<td>67 (42.9)</td>
</tr>
<tr>
<td>Very good</td>
<td>36 (23.1)</td>
</tr>
<tr>
<td>Excellent</td>
<td>4 (2.6)</td>
</tr>
<tr>
<td>Personal loss/misfortune in last year, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes, one</td>
<td>32 (20.5)</td>
</tr>
<tr>
<td>Yes, more than one</td>
<td>23 (14.7)</td>
</tr>
<tr>
<td>No</td>
<td>101 (64.7)</td>
</tr>
<tr>
<td>Taken care of someone seriously ill, n (%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>103 (66)</td>
</tr>
<tr>
<td>No</td>
<td>53 (34)</td>
</tr>
<tr>
<td>Trust in the health care system, n (%)</td>
<td></td>
</tr>
<tr>
<td>Completely distrust</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Somewhat distrust</td>
<td>16 (9.9)</td>
</tr>
<tr>
<td>Neither</td>
<td>23 (14.3)</td>
</tr>
<tr>
<td>Somewhat trust</td>
<td>58 (36)</td>
</tr>
<tr>
<td>Completely trust</td>
<td>14 (8.7)</td>
</tr>
<tr>
<td>No response</td>
<td>50 (31.1)</td>
</tr>
</tbody>
</table>

Reach, Adoption, and Usability of the Digital Koda Platform

Of the 2850 patients who were sent the invitation to participate through their EHR patient portal, 183 participants read the message and clicked on the Koda link to enroll (6.4% response rate). Of those who responded to the invitation, 88% (n=161) created a Koda Health account to begin the study. Of the 161 participants who began their care plan on the platform, 129 (80.1%) completed the Koda platform intervention (Figure 2).
As measured by the SUS, the participant-reported usability of the digital Koda platform was 76.2, indicating good system usability (Figure 3). Out of possible scores from 0 to 100, the SUS scores ranged from 47.5 to 100.
Figure 3. Demonstration of the acceptability and score range interpretations for the Koda platform. The score was 76.2 on the System Usability Scale, indicating that the platform was acceptable and had a good score.

Knowledge of ACP

Wilcoxon signed-rank tests showed that there was a significant increase in knowledge of ACP after using the Koda platform compared to baseline. Ratings increased on all 4 knowledge items: knowing what ACP is ($Z = 119.5; P < .001$), knowing who can be an SDM ($Z = 202; P < .001$), what makes a good SDM ($Z = 235; P < .001$), and the amount of flexibility an SDM can have ($Z = 289; P < .001$). Table 2 shows means at each time point.
Table. Change in knowledge of advance care planning (ACP) before and after completing the digital platform (on a 5-point scale).

<table>
<thead>
<tr>
<th>Knowledge items</th>
<th>Before Koda use, mean rating (SD)</th>
<th>After Koda use, mean rating (SD)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knowing what ACP is</td>
<td>3.23 (1.31)</td>
<td>4.19 (0.97)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Knowing who can be an SDM</td>
<td>3.63 (1.14)</td>
<td>4.28 (0.93)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Knowing what makes a good SDM</td>
<td>3.60 (1.16)</td>
<td>4.28 (0.90)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Knowing what flexibility an SDM can have</td>
<td>3.26 (1.29)</td>
<td>4.19 (0.96)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

aSDM: surrogate decision-maker.

Readiness to Make Decisions

Wilcoxon signed-rank tests showed that there was a statistically significant increase in readiness for ACP decisions after using the Koda platform compared to baseline. Ratings increased on 3 readiness items: comfort level thinking about serious illness care ($Z=117; P<.001$), readiness to discuss future medical care with an SDM ($Z=316.5; P=.01$), and readiness to discuss future medical care with a health care provider ($Z=150; P<.001$). The postassessment revealed a statistically significant decrease in ratings for how worried patients were about future illnesses ($Z=536.5; P<.001$). Table 3 shows means at each time point.

Table. Change in readiness for advance care planning (ACP) decisions and conversations before and after completing the digital platform (on a 5-point scale).

<table>
<thead>
<tr>
<th>Readiness items</th>
<th>Before Koda use, mean rating (SD)</th>
<th>After Koda use, mean rating (SD)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comfort level thinking about serious illness care</td>
<td>2.89 (0.88)</td>
<td>3.06 (0.88)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Ready to discuss ACP with an SDM</td>
<td>3.23 (1.46)</td>
<td>3.59 (1.32)</td>
<td>.007</td>
</tr>
<tr>
<td>Ready to discuss ACP with health care provider</td>
<td>2.78 (1.37)</td>
<td>3.26 (1.18)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Worried thinking about future illness</td>
<td>3.32 (1.03)</td>
<td>3.07 (1.03)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

aSDM: surrogate decision-maker.
bThis item was reverse scored, so lower scores indicate higher ACP readiness.

Discussion

This pilot study assessed the feasibility and acceptability of using a web-based interactive ACP platform, Koda Health, to help older adults think about their overall health-related goals, document those goals in an advance directive, and assign an SDM. We found that the intervention was acceptable to older adults and feasible to implement. This was highlighted by the fact that 80% of enrolled participants completed the entire intervention and created an advance directive.

Our findings also suggest that the Koda platform was able to bridge the gap in ACP engagement between different racial groups. Previous reports have shown that individuals from often-marginalized racial communities are less likely to have engaged in ACP [26,28]. However, when comparing the 2 main self-identified race categories in our data, we found no substantial difference—80.95% of Black patients and 79.66% of White patients who enrolled in the study went on to complete using the Koda platform. This adds to the existing literature that indicates that digital ACP platforms have the potential to address health disparities by providing accessible, user-friendly tools to all users, regardless of their racial background [43,55].

In addition, the SUS score of 76.2 indicates that users generally found the Koda platform easy to use and had a positive overall experience. This score suggests that Koda’s interface and features were well designed, allowing users to navigate and interact with the system without substantial barriers. We also found that patient age was not correlated with system usability ratings ($r=0.03; P=76$), suggesting that the platform was similarly user-friendly across the sample age range (50-99 years). However, while the overall usability was rated as good, there may still be room for improvement. It will be important to analyze usability metrics in more detail to identify specific areas where the platform can be enhanced. This could involve conducting further user testing, collecting qualitative feedback, or conducting additional surveys to gather more insights.

Participants also reported learning new information about ACP and being more ready to have conversations about medical care after completing the plan. While the self-reported changes in knowledge and readiness were statistically significant, we cannot yet accurately determine the degree of clinical significance. Generally, clinically significant improvements are associated with any positive increase to an average 5-point rating. However, more research is needed to determine specific thresholds for outcomes on ACP-related metrics [17,49]. Nevertheless, these findings have positive implications for the ability of online health tools to promote ACP participation among patients and families, with the ultimate goal of bringing empowerment and peace of mind during serious illness or end-of-life care.
Despite these promising findings, several limitations must be acknowledged. First, the study intervention necessitates further validation through a randomized controlled trial. Second, the low enrollment rate compared to the total number of invitees suggest that more effective recruitment methods are needed than a single patient portal message, such as personalized ACP information from a nurse or other health care provider, posters placed in prominent areas, or additional contact methods. Additionally, as in any voluntary response, responses may have been affected by self-selection bias; those who agreed to participate in the study may be distinct from those who chose to ignore the invitation to participate. We saw evidence of this in the greater percentage of female-identifying participants in the study than in the general population. Future studies should consider appropriate sampling techniques like stratified randomization to ensure participation reflective of the larger population. Lastly, further studies should investigate the long-term impacts of the Koda platform on measures of patient and caregiver experience and goal-concordant care [54].

In conclusion, the Koda ACP platform represents a promising tool for promoting patient engagement in ACP, particularly among older adults and marginalized groups. By facilitating knowledge acquisition and readiness to engage in ACP, the Koda platform can help empower patients to make goal-informed medical decisions, especially regarding end-of-life care. Further research is needed to validate these findings and determine long-term impacts on patient and caregiver outcomes.

Acknowledgments
This study was supported by a National Science Foundation grant (award 2051460) and the Wake Forest Clinical and Translational Science Awards (grant UL1TR001420). JG was supported by the National Institute on Aging/National Institutes of Health (award K23AG070234). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors’ Contributions
DPM, TYF, KDC, EK, JM, AM, AEB, and JG conceptualized this study. RLR, KDC, TS, and JG drafted the manuscript. JM, AM, TS, and JG contributed to the editing of the manuscript. All authors approved the final manuscript.

Conflicts of Interest
RLR, KDC, DPM, and TYF were each employed by Koda Health at the time of their work on this project. TS, EK, AM, JM, AEB, and JG have no conflicts of interest to report.

References


Abbreviations

**ACP:** advance care planning  
**AHWF:** Atrium Health Wake Forest Baptist  
**CPR:** cardiopulmonary resuscitation  
**EHR:** electronic health record  
**SDM:** surrogate decision-maker  
**SUS:** System Usability Scale
Performance Differences of a Touch-Based Serial Reaction Time Task in Healthy Older Participants and Older Participants With Cognitive Impairment on a Tablet: Experimental Study

Christian Mychajliw, Dipl-Psych; Heiko Holz, BSc, MSc, Dr rer nat; Nathalie Minuth, BSc, MSc; Kristina Dawidowsky, BSc, MSc; Gerhard Wilhelm Eschweiler, Dr med; Florian Gerhard Metzger, Dr med; Franz Wortha, BEng, MSc, Dr rer nat

1. Geriatric Center, University Hospital for Psychiatry and Psychotherapy, University of Tübingen, Tübingen, Germany
2. TuCAN, Tübingen Cognitive Assessment for Neuropsychiatric Disorders, Tübingen, Germany
3. Institute of Computer Science, Ludwigsburg University of Education, Ludwigsburg, Germany
4. LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany
5. Vitos Hospital for Psychiatry and Psychotherapy Haina, Klinik für Psychiatrie und Psychotherapie, Vitos Haina gGmbH, Haina, Germany
6. Centre for Early Mathematics Learning, School of Science, Loughborough University, Loughborough, United Kingdom

*these authors contributed equally

Corresponding Author:
Christian Mychajliw, Dipl-Psych
Geriatric Center
University Hospital for Psychiatry and Psychotherapy
University of Tübingen
Calwerstraße 14
Tübingen, 72076
Germany
Phone: 49 07071 ext 2983538
Email: christian.mychajliw@med.uni-tuebingen.de

Abstract

Background: Digital neuropsychological tools for diagnosing neurodegenerative diseases in the older population are becoming more relevant and widely adopted because of their diagnostic capabilities. In this context, explicit memory is mainly examined. The assessment of implicit memory occurs to a lesser extent. A common measure for this assessment is the serial reaction time task (SRTT).

Objective: This study aims to develop and empirically test a digital tablet–based SRTT in older participants with cognitive impairment (CoI) and healthy control (HC) participants. On the basis of the parameters of response accuracy, reaction time, and learning curve, we measure implicit learning and compare the HC and CoI groups.

Methods: A total of 45 individuals (n=27, 60% HCs and n=18, 40% participants with CoI—diagnosed by an interdisciplinary team) completed a tablet-based SRTT. They were presented with 4 blocks of stimuli in sequence and a fifth block that consisted of stimuli appearing in random order. Statistical and machine learning modeling approaches were used to investigate how healthy individuals and individuals with CoI differed in their task performance and implicit learning.

Results: Linear mixed-effects models showed that individuals with CoI had significantly higher error rates (b=-3.64, SE 0.86; z=-4.25; P<.001); higher reaction times (F[1,41]=22.32; P<.001); and lower implicit learning, measured via the response increase between sequence blocks and the random block (β=-0.34; SE 0.12; t=-2.81; P=.007). Furthermore, machine learning models based on these findings were able to reliably and accurately predict whether an individual was in the HC or CoI group, with an average prediction accuracy of 77.13% (95% CI 74.67%-81.33%).

Conclusions: Our results showed that the HC and CoI groups differed substantially in their performance in the SRTT. This highlights the promising potential of implicit learning paradigms in the detection of CoI. The short testing paradigm based on these results is easy to use in clinical practice.

(JMIR Aging 2024;7:e48265) doi:10.2196/48265
KEYWORDS
serial reaction time task; SRTT; implicit learning; mobile digital assessments; cognitive impairment; neurodegeneration; tablet-based testing; mild cognitive impairment; MCI; dementia; Alzheimer; neuropsychology; aging; older individuals

Introduction
Memory, Neurodegeneration, and Aging
Overview
In an aging society, the number of individuals with neurodegenerative diseases is increasing. Alzheimer disease (AD) and Parkinson disease rank among the most prevalent neurodegenerative disorders. One of the most apparent cognitive symptoms of neurodegenerative diseases is a change in memory impairment, which can affect different cognitive and memory functions in different ways.

Memory functions differ in concepts and models [1-5]. When comparing different memory models, memory can be categorized into specific subsystems: intentional learning leads to explicit memory, and unintentional and incidental learning creates implicit memory [3-6]. Most studies on neuropsychological changes in neurodegenerative diseases focus on explicit memory, whereas studies examining the decline of implicit memory remain scarce [7-13]. Accordingly, many assessments for explicit memory are available, but only a few assessments are available for implicit memory, and most assessments are paper based. Paper-based examination of implicit memory is effortful and limited in its possibilities and test quality, although there are some digital tools available [8,14]. Still, solely examining explicit memory seems to be insufficient, as neurodegenerative disorders such as AD show changes in different cognitive domains, not just explicit memory but also implicit memory [14-19], among others. Thus, implicit memory paradigms are promising tools in addition to common explicit memory tasks in early diagnostics to assess memory dysfunctions more precisely and to determine the different causes of neuropsychological dysfunction. Reasons for reduced performance in implicit memory tasks can be independent of the reasons for deficits in regular tasks such as word lists used for explicit learning.

We introduce a digital tablet–based version of the serial reaction time task (SRTT) aimed at assessing implicit memory. We then use data gathered by this tool to develop machine learning (ML) models for predicting cognitive impairment without relying on diagnoses from comprehensive assessments such as the Montreal Cognitive Assessment or the Consortium to Establish a Registry for Alzheimer’s Disease. On the basis of computer-based studies, we developed an app for clinical application, aiming to provide results comparable with previous findings on the SRTT. This study examines a mobile touch–based SRTT on a tablet in healthy older participants and older participants with cognitive impairment (CoI).

Implicit Learning
Researchers use various terms for implicit memory, referring to different concepts and processes inconsistently. Sometimes, these terms can refer to subtypes of a form of memory [8,14]. Although implicit learning is an umbrella term for the absence of awareness and intention, which means “a collection of abilities that are expressed through performance without requiring conscious memory content” [2], statistical learning “refers to the ability to detect and learn regularities in the environment” [13]. Sequence learning “describes the ability to incidentally acquire knowledge of sequences of events and actions” [20]. Motor skill learning “refers to the increasing spatial and temporal accuracy of movements with practice” [21]. Procedural learning is used as a synonym for sequence learning, referring to the learning of sensory-motor skills [10,13,22,23]. As we are focusing on a rather application-oriented approach, we further refer to the more general concept of implicit learning.

SRTT Paradigm
Different tasks have been used to assess implicit learning [8,14]. Among others, the SRTT was established as a widely used assessment. In the original version of the SRTT [24], participants react to stimuli presented in blocks with repeated sequences. After several blocks of repeated sequences (“sequence blocks”), a block with random sequences is presented (“random block”). Although reaction times usually improve throughout the blocks of repeated sequences, they decelerate in the blocks with a random sequence [16,24]. Implicit learning is assumed when there is a decrease in reaction time in blocks of repeated sequences and an increase in reaction time in the random block. Figure 1 shows an illustration based on the example of this study.
SRTT Variants
There are many variations in the SRTT paradigm and ambiguous findings on the effects and results of different clinical syndromes. Variations in paradigms can significantly influence the patterns of results observed. Different versions of the SRTT may vary in factors such as the quantity and length of stimuli; number of trials and blocks; the arrangement, structure, and display of stimuli on screen; the method of response; and even the medium used. Stimuli are presented on a computer screen in most studies [14,25]. Participants are asked to indicate the position of the stimuli using buttons on a keyboard or button box. Very few studies used a touchscreen- or tablet-based presentation of the SRTT [9,26-31]. Thus, researchers have many degrees of freedom in adjusting the paradigms of the SRTT, especially when implementing a touch-based version for older participants. In this setup, the participants use their fingers to indicate the position of the stimuli directly on the touchscreen. The most appropriate specifications may vary based on the research question (RQ) and the sample under investigation.

As described by Hong et al [16], an alternating design, that is, a design with an alternating sequence and random blocks [32-34], has advantages in distinguishing between motor and cognitive learning, but progressions in sequence tasks cannot be analyzed. Moreover, alternating sequences lead to a longer overall assessment time. In contrast to laboratory studies, design decisions are limited in a clinical approach. An SRTT version suitable for clinical use with older participants should be as short as possible to meet their stamina and motivation, especially in the case of CoI. Although some researchers suggest the superiority of alternating SRTT variants because of their capacity to discriminate between sequence-specific and general skill learning [35], we decided to use the SRTT in a tablet-based version as a short SRTT version that only needs 5 blocks and thus is much shorter and more usable in clinical contexts.

Possible Distinguishing Features
The patient and control groups differ for various outcome measures. First, reaction times can differ between groups in general, meaning that healthy controls (HC) are faster than patient subgroups. Second, learning curves can vary between groups, meaning that HC participants should learn sequences faster. Third, the response increase between the groups may deviate. That is, the contrast in reaction times between sequence and random blocks becomes more pronounced after extensive learning of sequences, rather than just motor skill learning or increased familiarity with the task. Finally, the number of correct responses, that is, the response accuracy (and vice versa error rates), is expected to differ between groups. That is, the CoI group should show more false responses than the HC group. In addition, when we combine these variables as features in a statistical model, we may discover findings not only about implicit memory but also about parameters such as limitations in task comprehension or altered reaction times, which serve as additional diagnostic information.

SRTT Findings
The SRTT and similar tasks to assess implicit learning have been used in numerous studies in different fields [25], reporting different variables, outcome measures, and results. Varying patterns of results can be explained by different design variations (eg, [32,33]) and experimental requirements and conditions [8,14]. In healthy adults, differences in response increase were found consistently between sequence and random blocks. There was a slight tendency of age-related deterioration in performance, learning, accuracy, and reaction times with higher age [20,36,37]. Worse performance can be observed in patients with strokes [38]. Numerous studies show a deterioration of performance in neurodegenerative disorders affecting the basal ganglia and the thalamus, such as Parkinson disease [25,39,40], and other neurodegenerative and neuropsychiatric diseases, such as Huntington disease [41-43] and Korsakoff syndrome [24].

Some SRTT studies were conducted with patients with AD as the clinical sample, whereas others were conducted with patients with mild cognitive impairment (MCI; for reviews, refer to the studies by de Wit et al [8] and van Halteren-van Tilborg et al [14]). Overall, there are mixed findings on the association between SRTT performance and cognitive impairment (such
as AD, which can be a later stage of MCI). Comparing various studies poses a challenge because of differences in sample characteristics, such as the varying degrees of impairment severity (eg, MCI vs AD at different levels of severity), and the diverse inclusion criteria used for clinical samples (eg, specific diagnostic criteria vs various screening scores) [8]. In addition, inconsistent methodological approaches, including variations in the quantity and duration of stimuli, sequences, and blocks, further complicate comparisons between studies. Some studies revealed no significant differences in learning concerning response increase for participants with MCI compared with HC participants. That is, participants with MCI and HC participants showed similar response increases between sequence and random blocks [9,16,44,45], whereas other studies found differences [15,46]. Participants with AD, however, show less response increase than HC participants in some studies [47-49] and a comparable response increase in others [50-52].

Compared with learning curves, that is, the reduction in reaction time over sequence blocks, patients with MCI seem to have similar curves as HCs in most studies [9,15,16,44,45], whereas some studies found differences [46,48]. Patients with AD show more deficits than HCs, as indicated by a flatter learning curve in some studies [49,53], but this was not clearly evident in most studies [47,50-52].

When comparing accuracy or error rate, in some cases, no differences are found between participants with AD [41,51,52] and MCI [44,46]. In other cases, participants with AD [47,48,50] and MCI [16,45] differ from the reference groups.

In most cases, patients with MCI [16,46] and AD [47,50-53] had slower overall reaction time [54]. These differences can be explained by motoric demands, the experimental design of the tasks, and differences in sample selection. In addition, a large number of dementia diseases may be mixed pictures of different subtypes of dementia [55,56]. This phenomenon may further contribute to the divergent findings. In a review and meta-analysis, de Wit et al [8] discuss the difficulty of participants with AD in understanding and remembering the test instructions of SRTT paradigms. Most studies found differences in response accuracy and reaction times (with participants with CoI being slower than HC participants). These differences suggest that the understanding and execution of the task play a significant role in classifying differences between healthy individuals and individuals with CoI.

This Study

We developed a mobile SRTT version suitable for testing implicit memory in a clinical routine (in contrast to an extensive laboratory assessment). Working with older patients in everyday clinical routines, we recognized the need for a short, understandable, and highly accepted digital assessment that medical professionals can use in point-of-care or bedside tests without requiring additional technical equipment. We also expect a short and tablet-based variant to address the difficulties in task understanding and remembering task instructions among patients with CoI reported by de Wit et al [8]. Using a tablet in neuropsychological testing has benefits in terms of the availability of new data sources and its applicability outside the laboratory [57,58]. Furthermore, digitalized testing enables tasks and measures that are impossible in pen-and-paper testing [57]. Using a touchscreen is a considerable relief for older patients, and particularly patients with CoI, compared with using a keyboard [59-62]. Furthermore, motivational effects must be considered when designing tasks for older participants [57,63]. Not only the design of the task but also the task parameters are essential: the length of the sequence and the frequency of repetitions should capture implicit learning but should not be unnecessarily prolonged. Using ML approaches that combine various parameters, we may predict participants’ cognitive status more accurately with less data than a traditional approach, which accommodates shorter assessments. Combining the SRTT with ML represents a more recent development in this research domain [9,10,16].

ML predictions based on these parameters can potentially lead to accurate predictions using fewer repetitions, which facilitate shorter assessments that are mandatory for acceptance in clinical outpatient and inpatient practice.

Research Questions

This study focuses on the following RQs, which are centered on the question of whether the findings for the computer-based version of the SRTT can be replicated and transferred to the tablet-based version of the SRTT used in this study.

We investigated the following RQs:

1. Do participants with CoI and HC participants differ significantly in response accuracy?
   - We expect participants with CoI to make more errors than HC participants.

2. Do HC participants and participants with CoI differ in their average reaction times during the learning phase?
   - We expect that participants with CoI are systematically slower than HC participants.

3. Do participants with CoI and HC participants differ in implicit learning? That is, is the response increase in the random block compared with the learning curve significantly lower for participants with CoI?
   - We expect that participants with CoI show less implicit learning than HC.

4. Do participants with CoI show a different learning curve during the learning phase than HC participants?
   - We expect HC participants to show a steeper learning curve than participants with CoI, which should show a flatter learning curve.

5. Can we reliably predict participants’ groups using an ML prediction model?
   - We expect to classify participants with an accuracy comparable with that of Hong et al [16]. That is, we expect the 80.9% found by Hong et al [16] to be within our 95% CI of prediction accuracy.
Methods

Participants
We recruited and tested 49 older participants at the Geriatric Center at the University Clinic for Psychiatry and Psychotherapy in Tübingen, Germany. A total of 2 participants discontinued the experiment. One participant had to be excluded owing to a low response accuracy of 40%, which indicates a failure to understand and complete the task appropriately. We later identified 1 participant with significantly prolonged reaction times as an outlier and had to exclude this participant. Within the scope of this project, we also collected data from 11 participants with depression, which we excluded from the analysis of this study because of the focus on neurodegeneration. The remaining 45 participants (26 female individuals), aged between 52 and 87 (mean 68.4, SD 9.82) years, consisted of 27 HC participants and 18 participants with CoI. A list of inclusion and exclusion criteria is presented in Textbox 1.

We based the allocation of groups on experienced physician examinations, confirmed by an interdisciplinary team (physicians, psychologists, specialized therapists, and nurses), as most participants were known to us as patients of our (day) hospital and their caregivers or relatives. We also recorded the participants’ educational level. Subsequently, we converted the educational levels into corresponding years representing the time typically taken to achieve them. The descriptive statistics of the demographics of the sample are provided in Table 1.

Textbox 1. Inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults aged ≥50 years</td>
<td>Unable to perform or a lack of understanding of the task requirements</td>
</tr>
<tr>
<td>Diagnosis of cognitive impairment (for patient group), confirmed by an interdisciplinary team</td>
<td>Visual impairment</td>
</tr>
<tr>
<td>Understanding and agreement of informed consent</td>
<td>Refusal or inability to give informed consent</td>
</tr>
<tr>
<td>Participation on a voluntary basis</td>
<td>Acute delirious or psychotic episode</td>
</tr>
<tr>
<td></td>
<td>Acute medical or physical conditions</td>
</tr>
</tbody>
</table>

Table 1. Demographic data of the 2 groups (HC\textsuperscript{a} and CoI\textsuperscript{b}; N=45).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>HC (n=27)</th>
<th>CoI (n=18)</th>
<th>Total (N=45)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>64.89 (9.66)</td>
<td>73.67 (7.62)</td>
<td>68.40 (9.82)</td>
<td>.002\textsuperscript{c}</td>
</tr>
<tr>
<td>Range</td>
<td>52-85</td>
<td>55-87</td>
<td>52-87</td>
<td></td>
</tr>
<tr>
<td>Education (y)</td>
<td></td>
<td></td>
<td></td>
<td>.52\textsuperscript{c}</td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>11.74 (3.21)</td>
<td>11.06 (3.80)</td>
<td>11.47 (3.43)</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>8-17</td>
<td>8-19</td>
<td>8-19</td>
<td></td>
</tr>
<tr>
<td>Sex, n (%)</td>
<td></td>
<td></td>
<td></td>
<td>.39\textsuperscript{d}</td>
</tr>
<tr>
<td>Female</td>
<td>17 (63)</td>
<td>9 (50)</td>
<td>26 (58)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>10 (37)</td>
<td>9 (50)</td>
<td>19 (42)</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a}HC: healthy controls.
\textsuperscript{b}CoI: participants with cognitive impairment.
\textsuperscript{c}Linear model ANOVA.
\textsuperscript{d}Pearson chi-square test.
Materials

Overview

We used a tablet-based variant of the SRTT (described in the subsequent sections) designed for this study to meet the needs of older participants. The experiments were performed on a “Samsung Galaxy Tab A (2016) with S Pen” tablet (model SM-P580, Samsung Electronics) with a screen size of 10.1 inches, running on Android 7.0. Tablets were positioned on the table horizontally and planar in front of the participants (Figure 2). Thus, the participants were able to rest their elbows on the table. The participants’ task was to repeatedly respond as quickly as possible to the target stimulus, whose position changed, with their fingers.

Figure 2. View of the setup and task in the app used for the study. Participants were allowed to rest their hands on the table.

SRTT Design

The experimental design of the SRTT used in this study is based on the study by Lum et al [64], as they used a shorter version with fewer trials than previous studies, which was necessary in working with older participants. Thus, following a 10-trial practice phase, the SRTT used in this study consisted of 5 blocks with 60 trials each. Blocks 1 to 4 are sequence blocks and contain the 10-item sequence, repeated 6 times per block. The repeated sequence 4-2-3-1-3-2-4-3-2-1 is based on the original design of Nissen and Bullemer [24] and was also used by Lum et al [64] and Lum and Kidd [65]. In the circular order we used, the top position (north) corresponds to 1, 2 corresponds to the right-hand position (east), 3 corresponds to the lowest position (south), and 4 corresponds to the left position (west; Figure 3).

When we developed the tablet-based variant, a circular order of stimuli [9,10,64,66-69] was chosen to ensure comparable spatial distances between stimuli on the screen (Figure 3). A horizontal arrangement of stimuli would lead to unequal distances between stimuli. A touchscreen-based version of the SRTT was used in only a few studies in general [26,29,31,70], specifically in samples consisting of older individuals [27]. To the best of our knowledge, apart from the study by Dominey et al [27], no study with older participants that exclusively used touchscreen versions of the SRTT on a larger sample was published. The motor skills required for responding on a tablet surface differ from those needed for pressing buttons on a response panel [29,31].

In block 5, based on the study by Lum et al [64], the stimulus appears in a pseudorandomized order. This order is based on 2 conditions: first, each stimulus appears as often as in the antecedent (learning) sequence blocks, and second, the probability of appearing at 1 of the 4 positions after its antecedent stimulus is the same as in the learning sequence. We precomputed 1 pseudorandomized sequence (Multimedia Appendix 1) and used the same pseudorandomized sequence for each participant. We did not inform participants about the given configuration. Using the difference between sequential and pseudorandomized trials yields a measure of skill acquisition from the SRTT that is specific and sensitive, as measuring implicit learning by only comparing improved reaction times in sequenced blocks is confounded by visuomotor association [71]. Multiple parameters captured through the app were used for statistical modeling to assess their predictive value, via the approach of Hong et al [16] using random forest classification.

After the stimulus appears at position 1, the probability that it will appear at either position 3 or position 4 is 50%, respectively. After appearing at position 2, it is equally likely that the stimulus appears next at positions 1, 2, or 4, corresponding to 33% for each position. After appearing at position 3, the probability for the stimulus to appear next is 33.33% for position 1 and 66.66% for position 2. After appearing at position 4, the probability for the stimulus to appear next at position 2 or position 3 is 50%, respectively. After a practice trial, the participants were advised to react to the stimulus, changing their position as quickly as possible throughout the blocks. The app recorded the reaction times and the correctness of the reactions.
**Task Implementation**

We developed the mobile touch-based SRTT in Unity 3D, version 2019.1.0f2 [72] as part of the TuCAN (Tübingen Cognitive Assessment for Neuropsychiatric Disorders) Project, which develops a tablet-based test battery app. In a first pilot study with university students, we showed that different user interface designs on the tablet are comparable and that no effects are attributable to the design. Moreover, we examined the usability and preferences of different designs with older participants in a second pilot study. The usability study is substantial for older participants to accommodate for possible low computer and tablet literacy and to ensure that an app is developed according to the needs of older participants [60,73,74]. In the preceding user tests and pilot studies described in this section, in which different designs were compared, we identified a circular compass design as the preferred design version for older participants (Figure 4). In this design, a compass dial is placed in the center of the screen and is surrounded by 4 circles. In the background, an ancient-looking map is depicted. A ship’s wheel, as the target stimulus, changes the positions between the 4 circles.

**Figure 3.** Circular order of the serial reaction time task.

**Statistical Analysis**

**Overview**

We performed analyses on the full sample of 45 participants. In addition, because of significant differences in age between the groups, all statistical analyses were rerun using an age-matched subsample of 36 participants (18 HCs and 18 participants with CoI). The pattern of the results remained identical, underlining the robustness of the findings to age differences. Therefore, we only report the results from the full sample in this study. The results of the matched samples are reported in the web supplement [75]. Similarly, the assumptions for all statistical models were checked. In case of assumption violation, we reran the analyses with robust models to ensure that the pattern of results remained identical. For the sake of readability, the results of the robust models are only reported in the web supplement [75].

**Analyses Software**

We conducted statistical analyses for RQs 1 to 4 using R software (R version 4.3.2, R Foundation for Statistical Computing) [76].

We used the `lme4` package (version 1.1-35) to fit (generalized) linear mixed-effect models [77]. The df and P values were calculated with the `lmerTest` package (version 3.1-3) [78] using Satterthwaite approximation for the denominator df. We
calculated the CIs for logistic regressions with the `broom.mixed` package (version 0.2.9.4) [79] using Wald approximation. We used the `ggplot2` package (version 3.4.4) [80] to create plots, the `stargazer` (version 5.2.3) [81], the `arsenal` (version 3.6.3) [82], and `tab_model` from the `sjplot` (version 2.8.15) [83,84] packages to create tables.

The criterion of statistical significance was set at Cronbach $\alpha=.05$. The raw data and R scripts detailing all analyses can be accessed in the web supplement [75].

**Data Cleaning**

For all analyses except for the accuracy analysis (RQ 1), we removed trials within blocks according to the following criteria in the following order: (1) the first trial of each block, as these trials succeed the fixation cross; (2) erroneous trials; (3) trials following erroneous trials; (4) trials with reaction times <200 ms; and (5) trials with reaction times deviating >2.5 SDs from the mean within a block, within participants. In total, we removed 5.36% (723/13,500) of trials.

**General Modeling Approach (RQ 1+RQ 3+RQ 4)**

Mixed-effect models will be hierarchically constructed from a full model (containing all fixed and random effects, including interactions) with a maximum random effects structure to model the with the best fit according to the Bayesian information criterion by removing the most complex fixed effects first (ie, interaction terms). If the complexity of the random effect structure is not supported by the data (ie, convergence issues), the random effects structure is reduced, similar to the fixed effects structure, by removing the most complex terms first.

**Accuracy (RQ 1)**

To contrast the difference in accuracy between the 2 groups, we ran generalized linear mixed-effects models using the logistic link function in all trials. Accuracy was computed as the proportion of correct trials to the total number of trials per participant per block. The models included group, block, their interaction, and age as fixed effects and random intercepts for the participant.

**Average Reaction Time (RQ 2)**

Analyses of covariance (ANCOVAs) were performed on the cleaned data (ie, correct trials only) to contrast the difference in the average reaction time across the learning phase, that is, excluding the random block. The average reaction time was computed per participant as the average of the median reaction time of each block. The fixed effects were group, with age, sex, and years of education as covariates.

On the basis of recommendations for good scientific practice for reporting ANCOVAs [85], an ANOVA comparing the average reaction times between groups was also performed, showcasing the impact of the covariates on the results.

**Implicit Learning (RQ 3) and Learning Curve (RQ 4)**

Linear mixed effect (LME) spline models with the last sequence block as the knot were performed on participants’ median reaction times per block to contrast the influence of group, block, response increase, sex, age, education, and the interaction between block and group and response increase and group. The model included random slopes for block and response increase for the participants. The linear and quadratic effects of blocks were tested for their contribution to the model fit. To this end, orthogonal polynomials were computed to encode the linear and quadratic effects of time. The response increase between the expected reaction time in the random block, based on the estimated learning curve during the learning phase in the sequence block, and the measured reaction time in the random block was coded as follows: We used dummy coding for the response increase between the expected reaction time in the random block, based on the estimated learning curve during the learning phase in the sequence blocks and the measured reaction time in the random block. That is, the dummy variable for response increase is set to 0 for blocks 1 to 4, and to 1 for the last and fifth block (“Response increase = 1 if random block, else 0”).

**Prediction Model: Classification of Group (RQ 5)**

In an exploratory step, we trained random forests to investigate how accurately the participants’ group (CoI vs HC) could be predicted. The input features were participants’ mean accuracy across blocks, participants’ mean reaction time across the learning phase (refer to RQ 2), age, and participants’ estimated learning curve and response increase. In addition, as features for learning curves (linear and quadratic effect of block) and implicit learning (response increase, coded as explained in the Implicit Learning (RQ 3) and Learning Curve (RQ 4) section, we extracted the predicted values of the LME model with the median of the $z$-transformed reaction time per participant per block as the response and learning curve and implicit learning as fixed effects. The model included random slopes for the learning curve and the response increase of the participants.

**ML and Prediction Model (RQ 5)**

ML approaches have shown promising results in predicting potential diagnoses and outcomes. These predictive models combine various parameters that were collected during the study. However, only 1 study by Hong et al [16] used an ML approach to predict participants’ cognitive status. Using a random forest approach [86], they achieved a prediction accuracy of 80.9%. Thus, using ML approaches to predict participants’ potential diagnoses rather than only examining group differences may improve the value of such tasks in cognitive assessments, enabling their broader use in populations of older individuals.

As explained previously, we used the random forest classifier to predict whether a participant belonged to the CoI group or the HC group. Repeated nested leave-one-out cross-validations were used to optimize hyperparameters and gain unbiased estimates of the model performance (eg, [86]). Specifically, each training data set from the initial leave-one-out cross-validation (outer cross-validation) was further split using a subsequent leave-one-out cross-validation (inner cross-validation). In the inner cross-validation, the number of trees per forest (range 10-100 in steps of 10), their maximum depth (range 1-7), and the minimum number of samples in each leaf (range 1-5) were optimized using grid search. Subsequently, to obtain an unbiased measure of accuracy, the best model from the inner cross-validation was used to predict the test set from the corresponding outer cross-validation. Finally, this procedure

https://aging.jmir.org/2024/1/e48265

Mychajliw et al
was repeated 15 times (ie, the same cross-validation procedure with varying random seeds) to account for random variations in the modeling procedure. These ML analyses (RQ 5) were conducted in Julia (version 1.9.3) [87] using the Machine Learning in Julia (MLJ, version 0.19.2) library [88].

**Ethical Considerations**

The study was approved by the ethics committee of the University Hospital of Tübingen (332/2016BO2). Participation was on a voluntary basis and after written informed consent and signature. Compliance with data protection and the implementation and evaluation were based on relevant regulations, guidelines, and protocols.

### Results

**Response Accuracy (RQ 1)**

The descriptive statistics for the average response accuracy per group are shown in Figure 5. The most suitable generalized linear mixed-effects model using the logistic link function to predict participants’ accuracy was obtained through hierarchical model comparisons, as outlined in the modeling approach in the General Modeling Approach (RQ 1+RQ 3+RQ 4) section. The final model contained block, group, and the interaction of block and group as fixed effects and a random intercept for participants. Hierarchical model analyses revealed that age did not contribute significantly to the model fit. The model revealed a significant main effect of group ($b=-3.64$, SE=0.86; $z=-4.25$; $P<.001$). We further found a significant interaction effect between block and group ($b=0.53$, SE=0.23; $z=2.33$; $P=.02$). The main effect of block was not significant ($b=-0.18$, SE=0.11; $z=-1.54$; $P=.12$). The models are listed in Table 2.

![Figure 5. Average response accuracy by group and block. Bars represent the SE of the mean. CoI: participants with cognitive impairment; HC: healthy controls.](https://aging.jmir.org/2024/1/e48265)
Table 2. Estimates for participants’ response accuracy for the generalized linear mixed-effects models obtained in the hierarchical modeling approach. The final model is provided in the second column.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Dependent variable: response accuracy</th>
<th>R0: full model containing all fixed and random effects, including interactions&lt;sup&gt;a&lt;/sup&gt;</th>
<th>R1: final model, without age as fixed effect&lt;sup&gt;b&lt;/sup&gt;</th>
<th>R2: model without age and the interaction between group and block&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Odds ratio (CI)</td>
<td>z value</td>
<td>P value</td>
<td>Odds ratio (CI)</td>
</tr>
<tr>
<td>Intercept</td>
<td>270.03 (2.02-36181.50)</td>
<td>2.24</td>
<td>.03</td>
<td>552.97 (238.66-1281.22)</td>
</tr>
<tr>
<td>Group</td>
<td>0.02 (0.004-0.15)</td>
<td>−4.02</td>
<td>&lt;.001</td>
<td>0.03 (0.005-0.14)</td>
</tr>
<tr>
<td>Block</td>
<td>0.84 (0.67-1.05)</td>
<td>−1.54</td>
<td>.12</td>
<td>0.84 (0.67-1.05)</td>
</tr>
<tr>
<td>Age (y)</td>
<td>1.01 (0.94-1.08)</td>
<td>0.29</td>
<td>.77</td>
<td>__d</td>
</tr>
<tr>
<td>Group×block</td>
<td>1.70 (1.09-2.66)</td>
<td>2.33</td>
<td>.02</td>
<td>1.70 (1.09-2.66)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Observations=225; Bayesian information criterion=420.48.

<sup>b</sup>Observations=225; Bayesian information criterion=415.16.

<sup>c</sup>Observations=225; Bayesian information criterion=416.91.

<sup>d</sup>Variables do not apply to a specific model.

**Average Reaction Times During the Learning Phase (RQ 2)**

The ANCOVA comparing participants’ average reaction time (computed as the average of the median reaction times per block) during the learning phase between groups while controlling for age and education revealed a significant difference in the mean reaction times between participants with CoI and HC participants ($F_{1,41}=22.32; P<.001$), with a large effect size of Cohen $d_{estimated}=1.61 (\eta^2_p=0.35)$. Participants with CoI were, on average, 198.57 (SE 42.03) ms slower than the HCs (during the learning phase). Furthermore, an ANOVA comparing mean reaction times between groups without covariates was conducted to test the robustness of the findings. The results showed a significant difference in reaction time between the participants with CoI and the HC groups ($F_{1,43}=37.02; P<.001$), indicating a robust effect. Descriptive statistics for the average reaction times during the learning phase are presented in Figure 6, and the results of the statistical analyses are provided in Table 3.

**Figure 6.** Descriptive statistics of the average reaction times per group and block. Bars represent the SE of the mean. (A) untransformed reaction times and (B) $z$-transformed reaction times. CoI: participants with cognitive impairment; HC: healthy controls.
Table 3. Estimates for participants’ average reaction time during the learning phase for the analysis of covariance (ANCOVA) and ANOVA models.

<table>
<thead>
<tr>
<th>Variables</th>
<th>ANCOVA&lt;sup&gt;a&lt;/sup&gt;</th>
<th></th>
<th></th>
<th></th>
<th>ANOVA&lt;sup&gt;b&lt;/sup&gt;</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b (SE)</td>
<td>F test</td>
<td>η&lt;sub&gt;p&lt;/sub&gt;&lt;sup&gt;2&lt;/sup&gt;</td>
<td>P</td>
<td>b (SE)</td>
<td>F test</td>
<td>η&lt;sub&gt;p&lt;/sub&gt;&lt;sup&gt;2&lt;/sup&gt;</td>
<td>P</td>
</tr>
<tr>
<td>Intercept</td>
<td>526.71 (164.70)</td>
<td>10.23 (1,41)</td>
<td><em>c</em></td>
<td>.003</td>
<td>729.52 (19.22)</td>
<td>1440.53 (1,43)</td>
<td><em>c</em></td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Group</td>
<td>198.57 (42.03)</td>
<td>22.32 (1,41)</td>
<td>0.35</td>
<td>&lt;.001</td>
<td>233.88 (38.44)</td>
<td>37.02 (1,43)</td>
<td>0.46</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age (y)</td>
<td>3.67 (2.12)</td>
<td>2.99 (1,41)</td>
<td>0.07</td>
<td>.09</td>
<td><strong>d</strong></td>
<td><strong>d</strong></td>
<td><strong>d</strong></td>
<td><strong>d</strong></td>
</tr>
<tr>
<td>Education (y)</td>
<td>-4.52 (5.48)</td>
<td>0.68 (1,41)</td>
<td>0.02</td>
<td>.41</td>
<td><strong>d</strong></td>
<td><strong>d</strong></td>
<td><strong>d</strong></td>
<td><strong>d</strong></td>
</tr>
</tbody>
</table>

<sup>a</sup>R<sup>2</sup>=0.51, Adjusted R<sup>2</sup>=0.47, F<sub>3,41</sub>=14.18; P<.001.
<sup>b</sup>R<sup>2</sup>=0.46, Adjusted R<sup>2</sup>=0.45, F<sub>1,43</sub>=37.02; P<.001.
<sup>c</sup>Not applicable.
<sup>d</sup>Variables do not apply to a specific model.

**Implicit Learning and Learning Curve (RQ 3 and RQ 4)**

The descriptive statistics of the average reaction times per group are displayed in Figure 6. As the final model predicting the raw reaction times obtained through the modeling approach outlined in the Statistical Analysis section in the Methods section resulted in nonnormally distributed residuals, we decided to use z-transformation of reaction times over the complete experiment per participant to reduce the effect of baseline differences in reaction times between individuals, which reduced skewness in the distribution of the residuals across participants. That is, we entered the median of the z-transformed reaction times per participant per block as the response into the LME. For models predicting the z-transformed reaction time, the visual inspection of the residual plot did not suggest a significant deviation from a normal distribution.

The final LME predicting z-transformed reaction times contained linear and quadratic terms of block, response increase, group, and the interaction between response increase and block, age, and education as fixed effects and random slopes for linear and quadratic terms of block as well as response increase of participants (formula: median reaction time [z-transformed] ~ time [linear]+ time [quadratic]+response increase+group+age+education in years+response increase: group+time [linear]+ time [quadratic]+response increase|participant). The hierarchical model analyses revealed that the interaction effect between the linear and quadratic effects of block and group did not significantly improve the model fit. The final model revealed a significant main effect of the linear effect of the block (β=−0.44, SE 0.09; t=−4.74; P<.001), a significant main effect of the quadratic effect of the block (β=0.46, SE 0.07; t=6.59; P<.001), a significant main effect of response increase (β=0.23, SE 0.11; t=2.21; P=.04), and a significant interaction effect of response increase and group (β=−0.34, SE 0.12; t=−2.81; P=.01). Regarding the interaction effect, participants with CoI had a significantly lower response increase between the random block and the last sequence block compared with HC participants. That is, there was a difference in response increase of z-transformed reaction times, obtained from the final model (β=−0.34, SE 0.12), and the difference in untransformed response increase between participants with Col (mean 26.83, SD 46.09 ms) and HCs (mean 45.37, SD 35.59 ms) was 18.54 ms. The effects of group, age, and education were not significant. The final model, along with the models investigated using the hierarchical modeling approach, is provided in Table 4.
Table 4. Estimates for participants’ z-transformed response increase of the linear mixed-effects models of the hierarchical modeling approach. The final model is listed in panel C.

<table>
<thead>
<tr>
<th>Variables</th>
<th>β (SE)</th>
<th>t test</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: R0: full model containing all fixed and random effects, including interactions</strong>&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>270.03 (-0.12)</td>
<td>-0.95</td>
<td>.34</td>
</tr>
<tr>
<td>Group</td>
<td>0.07 (0.06)</td>
<td>1.18</td>
<td>.24</td>
</tr>
<tr>
<td>Time (linear)</td>
<td>-0.43 (0.09)</td>
<td>-4.57</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time (quadratic)</td>
<td>0.47 (0.07)</td>
<td>6.53</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Response increase</td>
<td>0.22 (0.11)</td>
<td>1.97</td>
<td>.06</td>
</tr>
<tr>
<td>Age (y)</td>
<td>0.001 (0.002)</td>
<td>0.44</td>
<td>.67</td>
</tr>
<tr>
<td>Education (y)</td>
<td>-0.01 (0.004)</td>
<td>-1.63</td>
<td>.11</td>
</tr>
<tr>
<td>Group × response increase</td>
<td>-0.43 (0.23)</td>
<td>-1.90</td>
<td>.06</td>
</tr>
<tr>
<td>Group × time (linear)</td>
<td>0.02 (0.19)</td>
<td>0.12</td>
<td>.90</td>
</tr>
<tr>
<td>Group × time (quadratic)</td>
<td>0.08 (0.14)</td>
<td>0.57</td>
<td>.57</td>
</tr>
<tr>
<td><strong>Panel B: R1: model without the interaction between group and quadratic time</strong>&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.12 (0.12)</td>
<td>-0.97</td>
<td>.34</td>
</tr>
<tr>
<td>Group</td>
<td>0.05 (0.04)</td>
<td>1.08</td>
<td>.29</td>
</tr>
<tr>
<td>Time (linear)</td>
<td>-0.44 (0.09)</td>
<td>-4.65</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time (quadratic)</td>
<td>0.46 (0.07)</td>
<td>6.59</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Response increase</td>
<td>0.23 (0.11)</td>
<td>2.09</td>
<td>.04</td>
</tr>
<tr>
<td>Age (y)</td>
<td>0.001 (0.002)</td>
<td>0.43</td>
<td>.67</td>
</tr>
<tr>
<td>Education (y)</td>
<td>-0.01 (0.004)</td>
<td>-1.63</td>
<td>.11</td>
</tr>
<tr>
<td>Group × response increase</td>
<td>-0.33 (0.15)</td>
<td>-2.24</td>
<td>.03</td>
</tr>
<tr>
<td>Group × time (linear)</td>
<td>-0.03 (0.17)</td>
<td>0.15</td>
<td>.89</td>
</tr>
<tr>
<td><strong>Panel C: R2: final model without the interaction between group and quadratic and linear time</strong>&lt;sup&gt;c&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.12 (0.12)</td>
<td>-0.97</td>
<td>.34</td>
</tr>
<tr>
<td>Group</td>
<td>0.05 (0.04)</td>
<td>1.16</td>
<td>.25</td>
</tr>
<tr>
<td>Time (linear)</td>
<td>-0.44 (0.09)</td>
<td>-4.74</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time (quadratic)</td>
<td>0.46 (0.07)</td>
<td>6.59</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Response increase</td>
<td>0.23 (0.11)</td>
<td>2.09</td>
<td>.04</td>
</tr>
<tr>
<td>Age (y)</td>
<td>0.001 (0.002)</td>
<td>0.43</td>
<td>.67</td>
</tr>
<tr>
<td>Education (y)</td>
<td>-0.01 (0.004)</td>
<td>-1.63</td>
<td>.11</td>
</tr>
<tr>
<td>Group × response increase</td>
<td>-0.34 (0.12)</td>
<td>-2.81</td>
<td>.007</td>
</tr>
<tr>
<td><strong>Panel D: R3: model without the interaction between group and quadratic and linear time, and without the interaction between response increase and group</strong>&lt;sup&gt;d&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.13 (0.12)</td>
<td>-1.04</td>
<td>.30</td>
</tr>
<tr>
<td>Group</td>
<td>-0.03 (0.03)</td>
<td>-1.05</td>
<td>.30</td>
</tr>
<tr>
<td>Time (linear)</td>
<td>-0.44 (0.09)</td>
<td>-4.74</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time (quadratic)</td>
<td>0.46 (0.07)</td>
<td>6.59</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Response increase</td>
<td>0.27 (0.11)</td>
<td>2.33</td>
<td>.02</td>
</tr>
<tr>
<td>Age (y)</td>
<td>0.001 (0.002)</td>
<td>0.44</td>
<td>.66</td>
</tr>
<tr>
<td>Education (y)</td>
<td>-0.01 (0.004)</td>
<td>-1.63</td>
<td>.11</td>
</tr>
</tbody>
</table>

<sup>a</sup>Observations=225; Bayesian information criterion=150.74.
Prediction Model: Classification of Group (HC vs CoI; RQ 5)

Random forest classification predicting the group (CoI or HC) was computed using (1) standardized ordinal linear, (2) quadratic trends in reaction time for the sequence blocks, (3) standardized response increase, (4) age (years), (5) education (years), (6) response accuracy, and (7) average reaction time in milliseconds as features. To extract features 1 to 3, we refitted the LME from RQ 3 without the fixed effects of group, age, and education as well as the corresponding interaction terms. These models showed an average prediction accuracy of 77.13% (95% CI 74.67%-81.33%) across the repeated, nested leave-one-out cross-validation. The receiver operating characteristics curve is shown in Figure 7.

Figure 7. Receiver operating characteristic (ROC) curve for the prediction of cognitive impairment.

Discussion

Principal Findings

Overview

In this study, we presented and evaluated a digital tablet–based app featuring a variant of the SRTT to facilitate the diagnosis of implicit learning and memory and use it to predict the assignment of the diagnosis of CoI using an ML modeling approach. The app focuses on use in clinical routines and is based on computer-based studies and the findings of the SRTT. We evaluated our tablet-based SRTT with 27 HC participants and 18 older participants with mild to moderate CoI. We performed statistical analyses to evaluate the replicability and transfer of the results of previous (computer-based) SRTT studies with older participants with CoI to our tablet-based version of the SRTT. In addition, we deployed an ML modeling approach using a random forest classification to predict the participants’ group assignments (HC vs CoI). On the basis of the RQs outlined in the Introduction section, the results indicate that we were able to transfer the findings of previous studies to a tablet-based implementation of the SRTT in this study. We found the same significant performance differences between HC and CoI groups, and our ML modeling approach achieves promising results in predicting participants’ group assignments. In summary, our results indicate that the SRTT paradigm is transferable to (touch-based) tablet devices, and the results obtained with our app are comparable with previously published findings. The RQs and their findings are as follows:

1. Do participants with CoI and HC participants differ significantly in response accuracy? That is, do participants with CoI elicit more errors than HC participants?
   • Participants with CoI conducted, on average, significantly more errors per block than HC participants. We found no interaction between the block and the group.

2. Do participants with CoI and HC participants differ in average reaction times during the learning phase? That is, are participants with CoI systematically slower than HC participants?
   • Participants with CoI showed a significantly slower reaction time—on average, approximately 200 ms slower than HC participants (during the learning phase), with a large effect size of Cohen $d_{estimated}=1.61$ ($\eta_p^2=0.35$).

3. Do participants with CoI and HC participants differ in implicit learning? That is, is the response increase in the random block compared with the learning curve significantly lower?
   • Participants with CoI showed a significantly lower response increase than HC participants on $z$-transformed reaction times.

4. Do participants with CoI show a different learning curve during the learning phase than HC participants?
• No significant differences were observed between participants with CoI and HC participants in terms of linear, quadratic, or cubic learning curves.

5. Can we reliably predict participants’ groups using an ML prediction model?
• A random forest classification achieved an average prediction accuracy of 77.13%.

In this study, we used a touchscreen-based version of the SRTT. Thus, our results may differ from those of previous studies using keyboard or button box input because of the change in medium. We found the same effects in our tablet-based version previously found in comparable SRTT setups [8,27]. Thus, changing the medium does not significantly change the pattern of implicit learning in older participants with little prior knowledge of technology. The analysis of the parameters of the SRTT with models predicting participants’ groups allows us to make predictions about cognitive status and diagnoses with a relatively high accuracy. Even if only limited statements about isolated and pure implicit learning are possible [71,89], the app can be applied in daily clinical routines with older participants to collect diagnostic neuropsychological information.

**Response Accuracy (RQ 1)**

We found response accuracy to be lower in participants with CoI, in line with previous findings [8,14]. A low response accuracy may indicate difficulties in understanding and memorizing the task instruction, considering the hypothesis of difficulties with task comprehension [8,71,89]. In this study, we altogether excluded 4 participants from the analysis owing to difficulties with the task or attention and behavior difficulties or discontinuation of the examination. Among them, we excluded 2 participants who failed to complete the experiment, 1 participant because of low response accuracy, and 1 with exceptionally prolonged reaction times. This number is comparable with reported exclusions in other studies [8]. After exclusion, participants from the CoI group still had a percentage of correct trials of approximately 95%, compared with approximately 99% to 100% in the HC group. However, both numbers were still very high. Given the assumption of a strong influence of understanding the task instruction as a foremost parameter, as discussed in a recent review [8], we would suppose a lower response accuracy. Thus, the exclusion we made was in a manner that did not result in a systematic methodological error [8]. In summary, we were able to replicate previously published findings on a computer-based SRTT. At the same time, the results indicate that differences in response accuracy are not caused by task in comprehensibility or methodological errors, indicating the ecological reliability and applicability of our findings.

**Reaction Time (RQ 2)**

Similarly, we found differences in reaction times between the CoI and HC groups in the learning phase (Cohen \( d_{estimated} = 1.61; \eta_p^2 = 0.35 \); participants with CoI were, on average, approximately 200 ms slower). This finding follows most previous studies on reaction times in general [90] and in the SRTT in particular [14,16,91]. This finding can be explained by general RT differences in participants with mild and moderate CoI and AD owing to vigilance, cognitive, and psychomotor impairments [92], for example, caused by degeneration of the locus coeruleus [90,93].

**Response Increase (RQ 3)**

As a third factor, we found a significant difference in the response increase between the groups, represented by the difference between the estimated reaction time of the fitted learning curve and the measured reaction time in the random block. As shown in Figure 6, we observed a response increase for both groups. However, the CoI group showed a significantly lower response increase than the HC group, which again is in line with previous findings [8,14]. A lower response increase indicates a less sustainable learning of the sequence. That is, this indicates more than just motor learning and growing familiarity with the task [71]. The response increase in the classical SRTT paradigm has been used as a valuable and verified measure for (differences in) implicit learning [8,24]. Although more basic research scholars recommend a more complex paradigm, for example, with alternating sequences to differentiate different forms of learning more precisely [71,89], we opted for the straightforward approach of contrasting sequenced blocks with a random block to gain a sensitive and specific measure of skill learning for practical use in everyday clinical practice through a short and easy-to-perform task.

**Learning Curve (RQ 4)**

According to RQ 4, all participants showed improved reaction times across the 4 sequence blocks, indicating learning gains in both groups. These findings align with those of previous studies with comparable paradigms and samples [8,14]. With a more differentiated group division, we may find differences in the slope of the curve, which may indicate a more distinguished learning gain in the HC group. Motor learning and familiarity with such tasks certainly interfere with this finding. To what extent motor learning and familiarity with the task affect the learning curve cannot be differentiated at this point.

**Prediction Model: Classification of Group (RQ 5)**

The prediction model obtained through a random forest classification showed an accuracy of 77.13% in predicting the participants’ group (HC vs CoI) correctly. This performance is comparable with that of the study by Hong et al [16], who achieved an accuracy of 80.9% with a similar but lengthier version of the SRTT containing 4 learning and 4 random blocks of 48 trials each (384 trials total). In contrast, our version consisted of 5 blocks of 60 trials each, for a total of 300 trials. In addition, we achieved our results with a more robust ML approach using repeated nested cross-validation. Taken together, we achieved comparable accuracies using only the relatively short and straightforward SRTT paradigm combined with a robust random forest classification. This fact indicates the acceptable accuracy of the diagnosis classifications, despite only a coarse diagnosis classification. This insight is promising for future practical use.
Implicit Memory as a Part of Digital Neuropsychological Diagnostics

The development of neuropsychological deficits in explicit memory has been thoroughly researched and described and has become an integral part of dementia diagnostics. The role of implicit memory in the diagnosis and distinction of different subtypes of dementia has been scarcely investigated so far. In neuropsychological diagnostics and dementia research, implicit memory can be seen as an additional important domain in the entire pattern of deficits [94]. The use of digital assessment tools [57] can simplify examinations of implicit memory in clinical practice routines; even if overlaps in diseases exist, different participants show different deficit patterns in the process of neurodegeneration [95,96], partly also because of mixed subtypes of dementia [55,56]. This heterogeneity in neurodegeneration can also be seen as a relevant cause of inconsistent research outcomes [8,14] and, of course, needs further research on specific tasks such as the SRTT. On the basis of further research, differential diagnoses can be simplified using a tool similar to the one described in this study.

As different subtypes and mixed subtypes [55,56] have different progression types, paradigms such as the SRTT used for this study can help in the differential diagnosis of different dementia subtypes. When diagnosing the neuropsychological profiles of mixed dementia subtypes, a deeper and more differentiated examination at the level of explicit and implicit memory may be helpful. Using ML prediction can provide further benefit in differentiating diagnostic information based on future clinical studies that include more detailed and comprehensive diagnostics. Even this methodologically broad approach to group classification and the transdiagnostic and heterogeneous CoI group yielded significant results. Therefore, a more sophisticated approach to discriminate diagnosis groups will provide at least comparable results.

One goal of this study was to develop an assessment tool usable in clinical practice without exposing participants to unnecessary strains because of the length and complexity of the task. In developing a tablet-based tool relying on preceding user tests with older participants, we provided a short and transportable assessment instrument suitable even for older participants with CoI.

Relevance of This Study

In this study, we investigated a touch-based version of the SRTT in a sample of older participants. No control through hardware devices such as keyboards or response boxes was necessary; participants responded directly to the visual target stimuli with their fingers. The response increase and overall high response accuracy, even in participants with CoI, indicate that the paradigm we used is manageable and appropriate for older participants and that the original button-based paradigm is transferrable to tablets.

In our study, exploring an undifferentiated and roughly divided sample, significant differences between the groups were found. The application of statistical models enables the inclusion of features that exceed mere implicit memory, such as response increase. Therefore, the random forest trained achieved a prediction probability of the diagnosis groups of 77.13%. The accuracy of group prediction in our study is comparable with the accuracy reported by Hong et al [16].

Different definitions and concepts of learning are commonly used, based on different memory models, partly as different subtypes. We adhere to “implicit learning” as an umbrella term, as our task is too unspecific to distinguish more sophisticated terms and to differentiate which parts of the process can be explained through motor learning or sequence learning. Our essential objective was not to develop an experimental paradigm for the laboratory to distinguish forms of learning clearly but to provide a simple screening usable as part of a short battery of tests in clinical practice. Such tests could help to distinguish different diagnostic groups in real-life practice.

Limitations and Strengths

The study was initially part of a technical feasibility study for tablet use that did not address the conventional quality criteria of a clinical trial but had high ecological validity. Trained professional teams made the diagnoses after an extensive examination. Thus, classification into groups was based on the judgments of trained specialists and confirmed by interdisciplinary teams, as neuropsychological and depression scores were not available for all participants.

Divergent paradigms that can distinguish implicit and motor learning more sensitively are available. A more profound distinction between learning processes is not possible with the paradigm used in this study. We intended not to develop a tool for laboratory purposes but a user-centric tool that is usable in the clinic. Using statistical models, we are not limited to the exact distinction.

Taken together, the specific properties of the task partly explain the results found in our study. A pattern of stimuli alternating between random and sequence trials, for example, ensures the discrimination of explicit and implicit memory [32-34]. We chose a more focused approach without alternating patterns, as the use of statistical models for diagnostic information on implicit memory does not depend solely on accurately differentiated implicit memory processes in the experimental paradigm but on a variety of parameters. We also included parameters such as reaction times and learning gains in the statistical model. The objective of the task is not to map implicit learning as accurately as possible but to collect features that can be used in a statistical model to predict diagnostic information. The sequence length can be seen as a second factor. In this case, the well-proven and original sequence by Nissen and Bullemer [24] was used. Thus, we do not expect variations in the sequence or sequence length used in this study to result in the differences we found.

By contrast, a short tool that is easy to use in daily clinical practice is available to assess implicit memory on a tablet, for example, even at the bedside and not only in the laboratory. Because of the nature and implementation of the task, interruptions and early termination by participants who are stressed are less likely. In addition, participants do not receive negative feedback or feelings in the SRTT compared with explicit memory tasks, where they may experience failure in
repeating words, drawing figures, or calculating numbers. By not only including pure reaction times but also response increase and response accuracy in a ML model, reliable predictions regarding diagnoses can be made with relatively little data and within a short time. Comparable results, as in previous studies, can be achieved with our app more quickly and simply.

Overall, a reliable assignment of the diagnoses and high ecological validity are possible with the app’s relatively simple and short execution because of the use of ML algorithms. This assignment is preferable for a clinical setting, where brief assessments are essential. The data show that the short procedure is effective and yields results comparable with those obtained with more extended tests.

**Future Studies**

In the future, larger samples are needed to test the ability to discriminate similar conditions with heterogeneous cognitive symptom patterns such as dementia and delirium and different dementia subtypes. As implicit impairments are transdiagnostic, the SRTT and similar tasks have been examined with different samples [25]; however, only a few studies were performed with a touchscreen. Shortened versions of the task may facilitate the execution of the task to prevent cognitive overload in participants considered more impaired.

On the basis of more differentiated neuropsychological assessments, the parameters influencing test performance can be identified, especially when implementing additional information into statistical models. The SRTT can also be combined with another short task as another promising way to improve the accuracy of dementia diagnostics with the tablet. For example, this task could be a verbal task addressing different cognitive domains or a proven method such as the clock drawing test.

**Acknowledgments**

The study was performed as part of the TuCAN (Tübingen Cognitive Assessment for Neuropsychiatric Disorders) Project at the Geriatric Center Tübingen at the university clinic Tübingen, funded between 2018-2022 by the German Federal Ministry of Education and Research (BMBF, Start-interaktiv; grant 16SV8120). FW was supported by UKRI Economic and Social Research Council (grant number ES/W002914/1). The authors acknowledge the support from the Open Access Publication Fund of the University of Tübingen. The authors would like to thank all participants for their contribution.

**Authors’ Contributions**

NM, HH, and CM conceptualized the study. NM programmed the app and conducted the research as part of her master’s thesis under the supervision of CM, HH, FW, and KD. HH, CM, and FW performed the statistical evaluations. CM, HH, and FW wrote the first draft of the manuscript. FGM and GWE participated in the subsequent editing process. CM, HH, FW, FGM, and GWE read, reviewed, and edited the entire manuscript.

**Conflicts of Interest**

None declared.

**Multimedia Appendix 1**
Pseudorandomized sequence used in block 5.

[DOCX File, 15 KB - aging_v71le48265_app1.docx]

**References**


Performance Differences of a Touch-Based Serial Reaction Time Task in Healthy Older Participants and Older Participants With Cognitive Impairment on a Tablet: Experimental Study

Mychajliw C, Holz H, Minuth N, Davidsowsky K, Eschweiler GW, Metzger FG, Wortha F

Please cite as:

Mychajliw C, Holz H, Minuth N, Davidsowsky K, Eschweiler GW, Metzger FG, Wortha F
Performance Differences of a Touch-Based Serial Reaction Time Task in Healthy Older Participants and Older Participants With Cognitive Impairment on a Tablet: Experimental Study
JMIR Aging 2024;7:e48265
URL: https://aging.jmir.org/2024/1/e48265
doi:10.2196/48265
PMID:38512340

©Christian Mychajliw, Heiko Holz, Nathalie Minuth, Kristina Dawidowsky, Gerhard Wilhelm Eschweiler, Florian Gerhard Metzger, Franz Wortha. Originally published in JMIR Aging (https://aging.jmir.org), 21.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Smartwatch-Based Interventions for People With Dementia: User-Centered Design Approach

Doreen Goerss¹,², MD; Stefanie Köhler², MA; Eleonora Rong¹, MD; Anna Gesine Temp³, PhD; Ingo Kilimann¹,², MD; Gerald Bieber¹, PhD; Stefan Teipel¹,², Prof Dr

¹Department of Psychosomatic Medicine, Rostock University Medical Center, Rostock, Germany
²Deutsches Zentrum für Neurodegenerative Erkrankungen Rostock/Greifswald, Rostock, Germany
³Neurozentrum, Berufsgenossenschaftliches Klinikum Hamburg, Hamburg, Germany
⁴Fraunhofer Institut für Graphische Datenverarbeitung, Rostock, Germany

Abstract

Background: Assistive technologies can help people living with dementia maintain their everyday activities. Nevertheless, there is a gap between the potential and use of these materials. Involving future users may help close this gap, but the impact on people with dementia is unclear.

Objective: We aimed to determine if user-centered development of smartwatch-based interventions together with people with dementia is feasible. In addition, we evaluated the extent to which user feedback is plausible and therefore helpful for technological improvements.

Methods: We examined the interactions between smartwatches and people with dementia or people with mild cognitive impairment. All participants were prompted to complete 2 tasks (drinking water and a specific cognitive task). Prompts were triggered using a smartphone as a remote control and were repeated up to 3 times if participants failed to complete a task. Overall, 50% (20/40) of the participants received regular prompts, and 50% (20/40) received intensive audiovisual prompts to perform everyday tasks. Participants’ reactions were observed remotely via cameras. User feedback was captured via questionnaires, which included topics like usability, design, usefulness, and concerns. The internal consistency of the subscales was calculated. Plausibility was also checked using qualitative approaches.

Results: Participants noted their preferences for particular functions and improvements. Patients struggled with rating using the Likert scale; therefore, we assisted them with completing the questionnaire. Usability (mean 78 out of 100, SD 15.22) and usefulness (mean 9 out of 12) were rated high. The smartwatch design was appealing to most participants (31/40, 76%). Only a few participants (6/40, 15%) were concerned about using the watch. Better usability was associated with better cognition. The observed success and self-rated task comprehension were in agreement for most participants (32/40, 80%). In different qualitative analyses, participants’ responses were, in most cases, plausible. Only 8% (3/40) of the participants were completely unaware of their irregular task performance.

Conclusions: People with dementia can have positive experiences with smartwatches. Most people with dementia provided valuable information. Developing assistive technologies together with people with dementia can help to prioritize the future development of functional and nonfunctional features.

(JMIR Aging 2024;7:e50107) doi:10.2196/50107

KEYWORDS

assistive technology; user-centered design; usability; dementia; smartwatch; mobile phone
Introduction

Background

Dementia is associated with a loss of autonomy and restrictions in coping with everyday tasks [1], which often lead to caregiver burden [2]. There are still no curative treatments for dementia. Assistive technologies (ATs) can help people with dementia maintain their level of everyday activity [3]. To date, digital ATs have not been broadly applied in support and care for people with dementia.

Several systematic reviews regarding digital ATs for people with dementia [4-6] indicate increasing attention being given to wearable devices, for example, smartwatches, which represent the most intimate form of noninvasive ATs. Early digital ATs were aimed at increasing a person’s security by detecting falls and alerting caregivers. Current ATs interact with the wearer and address more than a single domain, for example, reminding the wearer about an event or detecting when the wearer falls [7,8]. The measurement of activities and physiological parameters and the application of user interfaces allow for more flexible support of daily living activities [7,9-12]. Despite the promising potential of ATs, many people with dementia do not use such technologies [13-15]. The reason for this could be that the needs of the target group were not adequately considered, for example, in terms of functional scope or usability of ATs [16]. According to the International Organization for Standardization and the International Electrotechnical Commission 9241-11 standard, usability is defined as the “extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use” [17].

User-centered innovations address unmet needs and play an important role in breaking barriers and increasing access to ATs [14,18]. Previous studies have shown that the analysis of stakeholder needs, wishes, and values is crucial for sustainable innovations [15,19,20], and focusing on users’ needs potentially prevents ATs from being nonusable or abandoned [13,21,22]. Considering the needs of future users from the beginning of development is mandatory from an ethical and a practical perspective [10,22].

However, placing people with dementia at the center of AT development can be challenging in the following ways: economically—patient involvement may increase the time and cost for organizations involved [23]; empirically—some scholars do not consider accounts of people with dementia to be reliable [24,25]; and ultimately, participation may distress or overwhelm people with dementia [25,26]. In addition, established tools for assessing user experience or usability may be insufficient and difficult to use for people with dementia because their ability to provide insight may be limited [27].

Objective

We analyzed the usability of a smartwatch application for addressing the needs of people with cognitive impairment based on a user-centered design approach together with people with mild cognitive impairment (MCI) or dementia. This study aimed to contribute to a better understanding of the values and limitations of user involvement in the development of a smartwatch to support people with dementia in their daily lives.

Methods

User-Centered Design Framework of the Sensor-Based Individualized Activity Management System for People With Dementia Study

The interdisciplinary Sensor-Based Individualized Activity Management System for People With Dementia (SAMi) study aimed to develop a mobile assistive device for people with memory impairments to support activities of daily living. The study was planned with a user-centered design approach from the beginning and included stepwise feedback from different stakeholders (Figure 1).

In step 1, we analyzed unmet needs. In a preparatory interview study, we conducted 30 semistructured interviews with stakeholders (people with dementia, health care professionals, and relatives of people with dementia) from a collaborating nursing home located in Pinnau and the geriatric ward of a hospital in Bad Doberan (both are small cities in Northeast Germany). We specifically asked about daily routines and situations in which people with dementia needed help or support. We also addressed the issue of technical assistance, design ideas, and circumstances that promote or hinder the acceptance and adoption of existing and potential future technologies. We applied qualitative content analysis according to Mayring [28] to analyze the material (refer to the study by Köhler et al [29]).

In a parallel, observational field study, we gave a passive smartwatch to 12 people with dementia residing in a nursing home and observed their behavior over a day. This smartwatch passively monitored activity via accelerometers and enabled indoor positioning via Bluetooth. It provided the users with no information except what time of day it was. During the monitoring period, participants’ behavior was annotated in real time by trained observers using the Pocket observer tool (version 3.3; Noldus IT), which included a customized annotation scheme (Multimedia Appendix 1). On the basis of the analysis of needs (step 1 in Figure 1), we conducted an intervention study that aimed to increase certain daily activities by prompting the participants. We have published the results of the interviews elsewhere [29]. The core element of this paper is the results of the intervention study, which represents the final part of the SAMi study.
Selection of Tasks for the Intervention Study
We decided to prompt participants to engage in 1 task each from the “activity” and “nursing” fields [29]. We implemented short instructions that completely appeared on the smartwatch’s screen. Consequently, we tested different tasks and task comprehension in person with patients from the memory clinic in step 2 (Figure 1). Finally, we implemented a prompt to drink some water (task A; “nursing”) and the “circling bells task” (task B; “activity”) on the smartwatch, as both tasks appeared to be comprehensible and suitable. Instructions for task B explicitly included the instructions to circle bells on a sheet. Further details about step 2 can be found in Multimedia Appendix 2.

Intervention Study

Study Population
The study is registered in ClinicalTrials.gov with the brief title, “SAIN_UMR” (NCT05885620). After a short test phase for system checks (step 3), we started the intervention study (step 4). Participants were recruited from the memory clinic of the Rostock University Medical Center. The inclusion criteria were being aged >50 years, having a diagnosis of MCI or dementia, and having a Mini Mental Status Examination (MMSE) score ≥9 and ≤28 points. The exclusion criteria were presence of clinically relevant impairment of visual acuity or hearing or relevant speech or language impairment. Patients were diagnosed according to international scientific diagnostic criteria, neuropsychological testing (Consortium to Establish a Registry for Alzheimer’s Disease [30]), and brain imaging; 78% (31/40) of the patients underwent cerebrospinal fluid analysis. Of the 40 patients, 12 (30%) were diagnosed with MCI [31,32], and 28 (70%) were diagnosed with dementia. Among these 28 patients, 24 (60%) were suspected to be in the Alzheimer spectrum: 18 (45%) with typical Alzheimer disease (AD) [33], 5 (13%) with mixed AD pathology [34,35], and 1 (3%) with atypical AD [33]. Of the 40 participants, 2 (5%) were diagnosed with primary progressive aphasia [36,37], 1 (3%) with a...
behavioral variant of frontotemporal degeneration [38], and 1 (3%) with Lewy body dementia [35].

Participants underwent clinical and neuropsychological examination—they received standardized examination of visual acuity. Near visual acuity was tested using a standardized optotype card with a decimal scale [39]. Distance acuity was assessed using a standardized eye chart at 4 to 5–m intervals (Oculus, number 4634). Both measurements were uncorrected and, when applicable, corrected with personal glasses. Hearing capabilities were assessed using a tablet-based certified app with a pure-tone threshold test (Mimi Health GmbH). Despite our efforts, we found that a standardized tablet hearing test could not be performed with our participants because they did not tap the button in time. Therefore, we decided to omit the regular test after 20 participants were assessed and did not include the results in our analysis.

As a global score for cognition, we used MMSE [40]. Visual constructive capabilities were assessed using a clock completion test [41] and the Rey complex figure direct copy (Rey Fig Copy) test [42]. Visual attention, processing speed, and task-switching abilities were measured using the Trail Making Test A (TMT-A) and Trail Making Test B (TMT-B) [43]. The results are presented in Table 1. We recruited 44 participants, 4 (9%) of whom had to be excluded. Of the 40 participants, 2 (5%) were diagnosed with subjective cognitive decline, and 1 (3%) had an MMSE score that did not match the inclusion criteria, and the trial procedure of 1 (3%) participant was incorrect because we missed repeating the intervention even though the participant did not complete the task. Finally, for the analysis, we obtained complete data sets from 40 participants. In summary, we included 50% (20/40) women and 50% (20/40) men with a mean age of 75 (SD 6.8; range 58-85) years.

Table 1. Demographics and characteristics of the participants. The table includes success scores based on observations and usability scores based on 10 items of the questionnaire.

<table>
<thead>
<tr>
<th>Age (y)</th>
<th>MMSE (points)</th>
<th>Rey Fig Copy (score)</th>
<th>CDT (Shulman score)</th>
<th>TMT-A (seconds)</th>
<th>TMT-B (seconds)</th>
<th>Visual acuity–near (score)</th>
<th>Visual acuity–distance (score)</th>
<th>Task A (score)</th>
<th>Task B (score)</th>
<th>Sum success score</th>
<th>Usability score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values, mean (SD)</td>
<td>74.98 (6.68)</td>
<td>23.70 (3.36)</td>
<td>23 (9.09)</td>
<td>2.55 (1.28)</td>
<td>99.10 (66.18)</td>
<td>184.30 (83.56)</td>
<td>0.46 (0.18)</td>
<td>0.69 (0.29)</td>
<td>0.80 (0.38)</td>
<td>0.65 (0.32)</td>
<td>1.45 (0.56)</td>
</tr>
<tr>
<td>Values, median (IQR)</td>
<td>77 (69-80)</td>
<td>25 (22-26)</td>
<td>25.50 (17.5-29)</td>
<td>3 (1-3)</td>
<td>72 (57.5-120.225)</td>
<td>174 (105-239)</td>
<td>0.45 (0.3475-0.5)</td>
<td>0.75 (0.5-0.83)</td>
<td>1 (1-1)</td>
<td>0.50 (0.5-1)</td>
<td>1.50 (1.5-2)</td>
</tr>
<tr>
<td>Minimum</td>
<td>58</td>
<td>12</td>
<td>0</td>
<td>1</td>
<td>31</td>
<td>52</td>
<td>0.20</td>
<td>0.13</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Maximum</td>
<td>85</td>
<td>28</td>
<td>35</td>
<td>6</td>
<td>329</td>
<td>384</td>
<td>1</td>
<td>1.66</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

aN=40; women: 20/40, 50%; men: 20/40, 50%; mild cognitive impairment diagnosis: 12/40, 30%; dementia diagnosis: 28/40, 70%; intense intervention mode: 20/40, 50%; regular intervention mode: 20/40, 50%.

bMMSE: Mini Mental Status Examination.
cRey Fig Copy: Rey complex figure direct copy.
dCDT: Clock Drawing Test (Shulman score).
eTMT-A: Trail Making Test A.
fTMT-B: Trail Making Test B.
gDrinking water.
hScore: 0=failure, 0.5=incomplete, and 1=completed.
iCircling bells.
jSum of task-A and task-B scores.

We conceptualized the observational intervention study based on the feedback obtained from the interview study and the experiences from the field study. Participants received interventions either in the “regular” or “intense” mode, under the observation of 2 cameras. Patients were assigned to one of the groups regardless of their neuropsychological test results. We applied an adaptive randomization procedure to balance the groups regarding participants’ age and sex. Compared to those in the “regular group,” the “intense group” received longer audio and vibration prompts and additional spoken output in response to the written instructions. Finally, a task-related picture appeared on both groups; this picture was animated in the “intense group.” To determine whether the order of the tasks had an effect on their completion, this was tested in both groups additionally. The final result was a study with 2 intervention arms (modus: intense or regular), each consisting of 2 subgroups (order of tasks: AB or BA; Figure 2).
To avoid anxiety, all participants were introduced to the study with a trial prompt initiated by a researcher immediately next to the participant. The trial prompt instructed participants to close their eyes. Then, the participants were instructed to make themselves comfortable and feel free to move within the room. When participants were familiarized with the device and the study procedure, the researchers left the room. After a 5-minute break, the first prompt was triggered remotely. All participants were given both tasks (drinking and circling bells), and prompts were repeated a maximum of 2 times if participants failed to comply. The time delay until repetition was set as 1 minute after the previous prompt.

Participants completed 2 questionnaires. One of the questionnaires captured participants’ previous experiences and affinity with technologies, and the other obtained feedback after wearing the smartwatch under camera observation. The summarized demographic information, test results, and outcomes are listed in Table 1, and additional details are provided in Multimedia Appendix 3.
**Technological Specifications**

A Huawei Watch 2 (4G) smartwatch was used. We designed the experiment in a Wizard-of-Oz-setting system using a smartphone as a remote control for the smartwatch (Figure 3). Consequently, we were able to repeat the interventions depending on the remotely observed participant’s compliance (success or failure) or to continue with the next intervention without needing instantaneous detection of behavior via sensors. We set specifications for the smartwatch based on previous experience [44,45] and updated our prototype iteratively; refer to steps 2 and 3 Figure 1. The smartwatch was set to a maximum brightness and volume, with a display size of 1.2 inches. Loudspeakers were limited to 85 dB due to European Union restrictions. In hearing distance, we measured the volume of audio output during interventions to be 56 to 66 dB. The volume varies during signals and speech output from the male voice. We remotely triggered video playback. This approach allowed full control of the font size and audio of the prompts. The length of the videos did not vary between groups. The smartwatch displayed time with the clock hands when no intervention was displayed.

Figure 3. Images and embedding of the applied prototype—smartphone with an app to manually trigger interventions on a watch, which is mounted on the patient’s wrist, showing instructions (in this case, “Bitte trinken Sie etwas Wasser,” which means “please drink some water”).

**Questionnaire**

We administered 2 questionnaires to the participants, 1 before and 1 after the intervention. The preintervention questionnaire contained three parts: (1) affinity for technology, based on the Affinity for Technology Interaction scale [46]; (2) personal experiences with technologies; and (3) motivation to participate. It contains 13 items. The postintervention questionnaire was developed according to the Technology Acceptance Model [47], System Usability Scale (SUS) [48], and Technology Usage Inventory [49]. The survey included six subscales: (1) usability, (2) design, (3) perceived usefulness, (4) concerns, (5) realization, and (6) experience during the study. A translated version of the questionnaire with items assigned to the subscales and item coding is available in Multimedia Appendix 4. This questionnaire contains 40 items: 33 closed questions to be answered on a 5-point Likert scale (from strongly agree to strongly disagree), 4 multiple-choice questions, 3 open-ended questions, and 1 opportunity for closing remarks by participants and staff. When we observed participants having difficulty with using the Likert scale, for example, due to alternating positively and negatively worded items, we moderated if necessary. In the moderating questionnaires, we assisted all the respondents while they were completing the questionnaires. If a participant was unsure about what the response scale indicated and whether “agree” or “disagree” indicated their opinion about the particular item, we explained the item in more detail. We also asked participants to review their answers when they accidently skipped questions.

**Internal Consistency, Usability Score, and Perceived Usefulness**

Overall, 3 items of the usability subscale were adapted from the SUS [48], whereas 7 were customized. To obtain a more intuitive score for the usability measurement, we processed the results of our 10-item usability scale analogous to the SUS [48]. In other words, each item was rated 0, 1, 2, 3, or 4 according to the answer on the Likert scale in the following direction: 0 was used for the strongest disagreement and 4 was used for the strongest agreement. The sum of all the scores was subsequently multiplied by 2.5, leading to possible usability scores ranging from 0 to 100. Higher scores indicate better usability.

Answers to the “perceived usefulness” subscale were interpreted similarly. We applied the scale from 0 to 4 for each of the 3 items, resulting in a sum of 0 to 12 responses per participant. Then, the average value of all participants was converted into percentage.

For the subscales related to usability, design, usefulness, and concerns, we calculated the internal consistency as Cronbach $\alpha$ and McDonald $\omega$ based on principal factor analysis using Jeffrey’s Amazing Statistics Program (JASP; version 0.16; JASP Team 2021; University of Amsterdam). Missing values were excluded pairwise. The reliability ($\alpha$ and $\omega$) ranges between 0 and 1. Higher values indicate greater agreement among items.
and suggest that participants’ responses were consistent. Cronbach $\alpha$ is a special case of McDonald $\omega$, whereas Cronbach $\alpha$ is based on the assumptions of unidimensionality, equal factor loadings, and uncorrelated errors, whereas McDonald $\omega$ accounts for varying factor loadings and error variances, making $\omega$ more appropriate to use. Cronbach $\alpha$ also is reported to be consistent with most previous literature.

**Measure of Success**

We observed reactions to the manually triggered interventions via video cameras and rated behavior based on a protocol to decide whether to repeat the intervention. Task A was rated as successful and scored 1 point if the participant drank some water. Task B was rated as successful and scored 1 point if the participant circled something on the worksheet with the pencil. Actions resulting in incomplete task fulfillment were rated with 0.5 points, for example, when a participant went to the table with the worksheet but without using the pen. When no activity that could lead to task fulfillment was initiated, 0 points were assigned. Only the best performance for each task was rated. Repetitions were not scored. The scores for both tasks were summarized, resulting in success values ranging from 0 to 2 points.

**Statistical Analysis**

Descriptive statistics, correlations, and 2-tailed $t$ tests were performed using JASP (version 0.16). To check for normality, we used the Shapiro-Wilk test. $P$ value >.05 was considered to indicate normal distribution. $P$ value >.05 in Leven test was considered to be consistent with the equality of variances. For variables that were not normally distributed, we calculated the Spearman rank-order correlation. Otherwise, we used Pearson correlation coefficient. For analysis of the questionnaire items, we used Kendall $\tau$, as the answers on the Likert scale are ordinal data. We chose a significance level of .05, with a corresponding confidence level of 95%.

**Ethical Considerations**

Ethics approval was obtained from Rostock Ethics Committee (A 2020-0071). All participants provided written informed consent.

**Results**

Demographics, clinical and neuropsychological test results, measures of success, and descriptive measures of the postintervention questionnaire are presented in Table 1 and Multimedia Appendix 5.

**Contributions From Respondents**

**Results Derived From Single-Choice Items**

Participants agreed very often with positively worded usability items (Multimedia Appendix 5). Most strongly agreed that they could sense vibration, hear sounds, and recognize visual cues well and had enough time to process the input. Of the 40 participants, 29 (73%) did not find the prompt disruptive. Of the 40 participants, 10 (25%) could not sense the vibration well. Interestingly, of the 40 participants, 23 (58%) agreed with the need to be supported by a technical person. Only one-third (14/40, 35%) were interested in more interactions with the smartwatch. When we processed the usability items analogously to the SUS [48], the mean usability score reached 78.3 (SD 15.4; range 30-100). The score translates as “good” usability score on the original SUS. The score did not significantly differ between men (mean 76.1) and women (mean 80.4; $P=.39$) according to the $t$ test. Usability decreased with age, but the effect was not statistically significant (Figure 4).

![Figure 4](https://aging.jmir.org/2024/1/e50107)  
**Figure 4.** Correlations and regressions between usability and age. The blue lines indicate CIs.

The design features were satisfactory to many participants. Most (30/40, 75%) found the appearance appealing, and 70% (28/40) found the watch to be properly sized. Further feedback about size was uniformly negative (12/40, 30%), with the watch being criticized for being very large. Of the 40 participants, 37 (93%) stated that the wristband was comfortable to wear. Some
participants said that the wristband should be softer (user10), longer (user23), or made from a different material (user17). It was suggested that its clasp be replaced with a magnet clasp (user11). Regarding usefulness, most participants (27/40, 68%) were interested in using the smartwatch frequently and saw a personal benefit in doing so (31/40, 78%). Approximately one-third (13/40, 33%) did not have any idea how the watch could be beneficial. Regarding perceived usefulness of the tool, participants scored an average value of 8.7 (corresponds to 72%), a median value of 9.5 (corresponds to 79%, SD 3.4) Only few were concerned that other people might hear the smartwatches’ announcements (17/40, 43%) or find the watch to be very conspicuous (12/40, 30%). Of the 40 participants, 6 (15%) were concerned about stigmatization because of the watch. Overall, two-thirds of the participants (27/40, 68%) felt a sense of safety while wearing the watch. Of the 40 participants, 5 (13%) were concerned about data protection. Of the 40 participants, 4 (10%) worried about their personal health when using the smartwatch. None of them (0/40, 0%) provided further explanations about these worries. Of the 40 participants, 29 (73%) expressed the wish for a permanent contact person for assistance ranged from “never” (user05) to “for all activities” (user29).

After completing the formal part of the questionnaire, the study staff asked participants who did not respond successfully to the prompts to comment individually. The responses revealed problems with different aspects of the drinking task. User02 said they did not recognize that the watch gave them this instruction. User17 initially could not find the water bottle. User05 did not dare to drink the water provided, and user23 answered that they were not thirsty. We also asked participants for further explanation when they circled the bell on the smartwatch screen instead of the worksheet. Of the 40 participants, 2 (5%) did not think they were supposed to do something in the real world (user33 and user05). User17 indicated that they did not realize that the instructions included the specification, “circle the bells on the sheet.” User34 said they heard the instructions but could not explain why they did not follow it. In addition, user10 stated that they had no idea what to do with the worksheet.

Internal Consistency of Subscales
We hypothesized that (1) usability, (2) design factors, (3) perceived usefulness, and (4) concerns play major roles in the user experience. Therefore, we organized the questionnaires into 4 subscales. Internal consistency was determined for each scale (Table 2).

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Items, n</th>
<th>McDonald ω (in descending order)</th>
<th>Cronbach α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Usefulness</td>
<td>3</td>
<td>0.81</td>
<td>.81</td>
</tr>
<tr>
<td>Concerns</td>
<td>7</td>
<td>0.79</td>
<td>.77</td>
</tr>
<tr>
<td>Design</td>
<td>6</td>
<td>0.75</td>
<td>.64</td>
</tr>
<tr>
<td>Usability</td>
<td>10</td>
<td>0.65</td>
<td>.64</td>
</tr>
</tbody>
</table>

Table 2. Internal consistency of subscales.
“Good” internal consistency ($\omega>0.8$) was achieved by the perceived usefulness scale, with concerns and design issues achieving “acceptable” consistency ($\omega>0.7$) and usability features achieving “questionable” levels of internal consistency ($\omega<0.7$). Overall, 3 items proved to be critical ($\Delta \omega>0.1$ if the item was excluded) for the internal consistency of their respective scales. Specifically, the items asked about the appropriate watch size, length of the instruction displayed, and desire for frequent use.

**Plausibility of Statements**

**Approach**

We evaluated the plausibility of the statements to examine the reliability of the feedback. In this paper, the term, “plausibility,” means “reasonable” or “consistent.” The responses to the items on our questionnaires were neither “right” or “wrong” nor expected by the researchers. Our approach to evaluating plausibility included 3 steps. First, we checked for discrepancies within the questionnaires; second, we compared the questionnaire responses and clinical assessments; and third, we compared the answers from the questionnaires with the observed task performance.

**Discrepancies Within the Questionnaires**

On an intrapersonal level, we found very few inconsistencies due to contradictory answers to different items. In the first step, we checked whether positively or negatively worded items were answered consistently, that is, whether the respondent agreed with the positive items and disagreed with the negative items and vice versa. Of the 40 participants, 6 (15%) agreed (“agree” and “strongly agree”) disproportionately with all items ($\geq$mean + SD). Of those 6 participants, 5 (83%; user01, user07, user08, user11, and user28) mostly disagreed with the 8 negatively worded items, indicating a positive attitude toward the smartwatch rather than a bias toward positive answers. Only 1 respondent (1/40, 3%; user03) who agreed with the positively worded items also agreed with the negatively worded items. She exhibited an uncritical tendency to confirm statements presented to her (“acquiescence bias”), making it unclear whether her answers truly reflected her point of view or if she misunderstood the questions.

In the second step, we analyzed the participants’ answers regarding the content of their statements. The most implausible answers were found in the questionnaire of user05. She personally completed the questionnaire and indicated high usability (usability score=90) and satisfaction with the design of our smartwatch. Otherwise, she disagreed with the desire to wear the watch in daily life and disagreed with all the questions about usefulness, although she already used a smartwatch in her daily life. She used open-ended questions, for example, to suggest a smaller size but did not explain the reason for indicating high usability and satisfaction with the app design despite her concerns. She answered the question for desired domains of support with “not at all” after canceling her initial answer, “always.” Fewer inconsistencies were found in the questionnaires of user26. She preferred single vibrations and short ringtones instead of multiple or repeated vibrations and longer ringtones, but recommended prolonged vibration and anticipated possible difficulties in hearing the sounds of the smartwatch for people with hearing impairment. In addition, user22 did not clearly indicate his preference regarding sounds. He selected both sound options even though they were mutually exclusive.

At a group level, it appeared implausible that less obtrusive notifications regarding audio (24>9) and vibration patterns (20>11) were favored, whereas speech output was a desired characteristic for most participants (24/40, 60%).

**Correlations Between the Questionnaire and Assessment Data**

When asked to mention situations in their daily lives that should be supported by a smartwatch, 11% (3/28) of the people with dementia stated no need for support (refer to the Items to Assess Practical Implementation section), although their caregivers reported a need for help that led to the diagnosis of dementia. This implies that these participants have lost insight into their functioning in daily life, limiting their reliability in naming domains for necessary support. Otherwise, they also anticipated a decline, as indicated by the added terms, “not yet,” “not to date,” and “not at present.” In addition, there was a discrepancy between the need for support for people with dementia and the severity of their impairment. One of the patients with MCI (1/12, 8%; user29) felt the need for support “for all activities.”

We also correlated the neuropsychological test results with the results from the subscales of the questionnaire. A better usability score (the higher the score is, the more usable the smartwatch appears to be) was significantly correlated with better cognition, as suggested by a positive association with the MMSE score ($P=.04$) and negative associations with the Shulman score on the clock completion test ($P=.02$) and time on the TMT-A ($P=0.11$) and TMT-B ($P=0.04$; Figure 5).

This confirmed our hypothesis that neuropsychological performance is associated with usability. The 3 other subscales did not correlate significantly with any neuropsychological test. Regarding sex, the 4 subscales did not significantly differ between the 2 groups.

We also checked for conformity of single items with the clinical test results. We correlated item 5 in scale 1 (recognizability of visual prompts on the smartwatch screen) with the visual acuity test results and found no relevant correlation ($\tau_{\text{Visual acuity}}=-0.24, P=.09$; $\tau_{\text{Visual acuity}}=-0.16, P=.23$) only. Only 5% (2/40) of the participants disagreed with good visibility. Both had a visual acuity below average. The 2 items focusing on latency, asking whether participants had sufficient time to notice that the watch notified them and whether instructions were shown for an adequate time, were correlated with the results of the TMT-A and TMT-B. In addition, no correlations were found ($\rho<0.3, \tau<0.2, P>.05$, respectively).
Questionnaires and Observed Task Performance

At the group level, single vibrations and less intrusive audio were preferred (refer to the Items to Assess Practical Implementation section). Regarding completion rates, participants in the “intense” group were more often successful. Here, we can see a discrepancy between the desire for less intrusive signals and better outcomes in participants with more intense intrusiveness.

We correlated the success (for the score calculation, refer to the Measure of Success section) with the usability score (Figure 6). Here, we found a significant positive correlation between the 2 measures; however, the effect size was small to moderate ($r=0.27$). The higher the usability was, the more successful the participant was.

We also checked for concordance between self-evaluations and observed behaviors at the individual level by investigating the overlap between success and specific items from the questionnaire. We divided each data set into 2 groups, resulting in 4 clusters (Figure 7). Regarding self-reflection, we distinguished between participants who agreed (clusters 1 and 3) and those who disagreed (clusters 2 and 4) based on item 7 of the usability scale, which assessed task comprehension. Regarding success, we distinguished between participants who completed no more than 1 task (clusters 2 and 3) and others (clusters 1 and 4).

We could see the concordance of self-reflection and observation in most participants (cluster 1: 26/40, 65% and cluster 2: 6/40, 15%). Of those people with appropriate self-reflection, most were successful (14/40, 35%) or almost completely successful (9/40, 23%). They correctly stated that they knew what they had to do. All the 12 people who were not completely successful had trouble with task B (circling bells). Of the 12 participants, 11 (92%) circled the bell on the smartwatch screen and 1 (8%) with the finger on the worksheet. In addition, participants from cluster 2, who were not able to complete >1 task, concordantly disagreed with good task comprehension. Interestingly, no one, including participants who did not attempt either task, strongly disagreed. Few (8/40, 20%) participants showed a discrepancy between self-reports and observations (clusters 3 and 4). We identified 3 possible reasons for the deviation in people in cluster 3. As shown previously, user05 was suspected to have had trouble in completing the questionnaire. We found inconsistencies within their questionnaire and suspected misinterpretation of questions or rating scales. Moreover, memory impairment might limit self-reflection. When completing the questionnaire, user02 indicated that she preferred to have had it handed out beforehand because she could not remember specific details (eg, vibration). User02 was diagnosed with AD dementia and had a score of 22 on the MMSE.

Finally, we found that incomplete or irregular fulfillment of tasks was not recognized by the concerned participants. Except for 1 participant (1/8, 13%; user02), all participants from clusters 3 and 4 completed task B and circled the bell image that appeared on the smartwatch screen. The individuals in cluster 3 felt that their understanding of the task was good, as they felt that they solved the task, even if they did not perform the task according to the researchers’ success protocol. This clear discrepancy between self-reflection and observation might indicate a loss of insight into their abilities and behavior. User33 explained that he did not consider using objects from the real world. User33 and user05 did not drink when prompted. Both belong to cluster 3. It is conceivable that some participants did not expect any actions involving their environment. Otherwise, all people in cluster 4 (5/5, 100%) managed the drinking task and circled the bell on the screen. It is unclear why they disagreed with task comprehension.
Summary
This paper describes an interdisciplinary study to identify the needs regarding and analyze the effects of multimodal interventions using auditory, haptic, and visual information provided by a smartwatch.

We hypothesized that perceived usability, design, usefulness, and concerns would influence the user experience with our
smartwatch and, consequently, included those topics using our questionnaire. We saw interest and commitment from the participants. Some participants struggled with the Likert scale, which might be avoided by guiding them or modifying the questionnaire. Usability was evaluated using complementary approaches, including questionnaire-derived measures and observational ratings of success based on completion rates. Usability, quantified through a 10-item score, reached a mean of 78 (maximum possible score=100). This finding is consistent with the remote observations: 35% (14/40) of the participants were able to complete both tasks, and 50% (20/40) of the participants solved one of both tasks. Only 15% (6/40) of the participants were unable to complete either task. Nonetheless, more than half of the respondents (23/40, 58%) thought they would need the support of a tech-savvy person to use the watch in the long term, and many (29/40, 73%) expressed the wish for a permanent contact person for maintenance. Only one-third (14/40, 35%) wanted more interaction with the smartwatch. The design was appealing to most participants, even though feedback about the size of the watch was primarily negative, with participants expressing that they felt the watch was very large. Various elements of the hardware have received suggestions for improvement. Perceived usefulness was rated high. Overall, two-thirds of the participants (27/40, 68%) had a sense of safety with the watch, and one-third of the participants (13/40, 33%) did not believe that the smartwatch would be beneficial. Only few respondents were concerned about using the watch. We also assessed nonfunctional requirements. We observed a preference for less intrusive notifications on the one hand and a desire for voice output on the other. Sophisticated feedback was provided about the wristband and the display latencies. Several specific use cases for possible reminder functions were mentioned. The internal consistency of the subscales varied between “questionable” and “good.” Lower cognitive ability was associated with lower usability.

We evaluated the plausibility of the feedback by checking for discrepancies. Only very few intraindividual inconsistencies were found within the questionnaires. There were no signs of a general acquiescence bias. When comparing self-rated task comprehension and observations, we observed concordance in 80% (32/40) of the participants. Mild deviation of self-reflection and task completion was observed in 13% (5/40) of the participants, and strong deviation was observed in 8% (3/40) of the participants. Participants seemed to lack insight into incomplete or irregular task performance. At the group level, we observed conflicting findings related to intrusiveness. Preferences for less audio and vibration conflicted with the desire for voice output, and more intrusiveness led to more successful task completion.

Discussion

Indications of Bias

Participants provided very favorable feedback. It is possible that our participants were not overly critical because they had little experience with better-designed, equivalent technologies. Older adults have a limited understanding of the potential implications for their privacy [50]. High usability could also be a sign of recruitment bias. Our recruitment strategy was not biased toward people who are interested in technology, as we invited all participants from the memory clinic. Previous experience with mobile technologies was not necessary. However, interest in the research topic is a major driver of participation; therefore, one must assume that participants were interested in technologies at least slightly.

We cannot rule out acquiescence bias, a tendency to choose the first option or a tendency to choose positive response options (agreement). We controlled for this bias by considering answers to inverted items: most of those items were consistently answered. This makes agreement bias unlikely. High levels of agreement possibly reflect a courtesy bias: the tendency to understate dissatisfaction or challenges with a system, driven by politeness. It is possible that respondents who anticipated positive feedback would be more desirable to researchers than those who provided negative feedback. At a metalevel, this poses a problem in participatory research and should be further investigated. When treated as co-researchers, people may be confused with different roles [51] and transform their commitment to science and innovations into less critical feedback in return for their empowerment. The participants were also reported to be motivated by family members and researchers [25]. Uncertainties regarding participants’ roles in participatory research have already been noted [26]. Otherwise, the participants who actively provide feedback or even participate as co-researchers are motivated to make a difference and transfer scientific results into practice [52]. Regarding ATs for people with dementia, professional researchers and co-researchers share the same need for technology transfer. Neither would benefit from spending resources for the development of unsuitable technologies as a result of uncritical evaluation. Transparency and clarification of expectations, roles, and goals could help resolve this conflict.

Rating of Usability and Review of Internal Consistency

Our calculated usability score cannot be compared directly with other scores resulting from SUS ratings, as we deviated from the original SUS by 7 items [48]. However, regarding the mean and SD, our score is similar to other SUS scores [53,54]. As described previously, usability decreases with age [53] and poorer cognition. We presented data about the internal consistency of the subscales of our questionnaire, with McDonald ω being between 0.65 and 0.81. These findings are consistent with those of previous studies about the internal consistency of usability questionnaires administered to older adults or people with cognitive impairment [54,55]; however, we found only a small number of studies regarding this topic.

Feasibility and Significance of User-Centered Design

There is an ongoing discussion about the extent to which people with dementia can be involved in research and in the development of ATs [20,25]. The ambiguous nature of participation has been examined in several recent reviews. Brett et al [23] showed many positive aspects of patient and public involvement in general; however, they also stressed that the outcomes were found more randomly than methodically. Kowe et al [26] summarized many advantages and disadvantages of participatory dementia research for researchers. Fischer et al
discovered positive and negative effects of involvement of older users in technology design but were unable to determine its impact on technology adoption and acceptance. However, Bethell et al [56] focused on engagement of patients with dementia but could not determine its impact on the research process or outcomes. All of them concluded that more evidence is needed to illustrate the impacts on the involved parties. Therefore, we aimed to determine the feasibility of user-centered development of smartwatch-based interventions for people with dementia.

We based our analysis on 3 aspects: study implementation, intervention outcome (success metrics), and qualitative measures (user experience and perceived usefulness). Regarding recruitment, we had no difficulty in identifying and enrolling suitable study participants. We found participants to be interested and dedicated to this field of research. We had only a small amount of missing data. This indicates that our study addressed a problem relevant to the target group. As we wanted to evaluate the feasibility of the user-centered design approach, we do not discuss about its technical feasibility in this paper.

Regarding intervention outcome, we could see that 80% (32/40) of the participants were able to completely solve task A and 40% (16/40) of the participants completely solved task B; a further 50% (20/40) of the participants completed task B at least partially. As these values are study specific and there are no values in the literature for comparison, it is difficult to determine what was to be expected. Many researchers portray older adults as technologically illiterate [20]. Many people in the target group had significant visual, hearing, and tactile impairments [57,58]; therefore, it was not clear whether a smartwatch would even be able to reach the threshold required to gain the attention and task understanding of these people. In this regard, the task completion rate in our study seems to be high. This high rate could also be explained by our interventions being common tasks requiring little effort. In addition, our tasks did not require participants to directly interact with the smartwatch.

Finally, regarding user experience, we received very positive feedback. Most participants felt comfortable with using the watch and would be willing to participate again in the study. Overall, three-fourth of the participants (31/40, 78%) could imagine having a personal benefit from using the watch, which is indicative of a high level of perceived usefulness. On the basis of the evaluation of these 3 aspects, we definitively consider the feasibility as given, but note that user involvement is more resource intensive [20].

The involvement of various stakeholders represents an additional expense [24]. On the one hand, it takes more time to conduct and, above all, evaluate interviews. On the other hand, the creation of suitable questionnaires is methodologically challenging. The scope (number of questions), the number of response levels, or the alternation of positively and negatively worded questions [59] are essential aspects that need to be considered [60]. The iterative approach [20] means that substeps and project goals cannot be defined and planned from the outset. Instead, interim evaluations and adjustments are necessary, which require consistent project management and effective team communication. In our case, the many years of expertise in this area and the commitment of the study staff contributed to our success, as did the long-term, third-party funded financial support over 3 years.

Plausibility Analyses

The second objective was to evaluate the extent to which user feedback from people with dementia regarding ATs is plausible. Our approach to evaluate plausibility included three steps: (1) analysis of discrepancies within questionnaires, (2) comparison of questionnaire responses and clinical assessments, and (3) comparison of questionnaires and observations.

Individual-based analyses revealed only isolated inconsistencies, and group-based analyses revealed inconsistencies that may indicate mutually exclusive needs. Participants showed a wish for less obtrusive notifications and dislike toward a watch that was very conspicuous but also stated a wish for speech output. This shows that there will be no technology that can address all needs equally. There can be conflicting needs both interindividually and intra-individually. This inconsistency in the statements is not due to cognitive deficits but reflects ambivalent attitudes that exist in all people to a greater or lesser extent [29]. We emphasize that unreliability in certain individual cases should not be a reason to prevent the whole group of patients with dementia from participating via self-reports [31].

Our results showed a moderate correlation between perceived usability and cognition (Figure 5). These findings were previously described [53] and confirmed our hypothesis that neuropsychological performance is associated with usability.

In contrast to previous studies [61], we observed high rates of overlap between self-reflection and success metrics such as completion rates. The concordance was evident in both directions. Objectively successful participants mostly indicated a good understanding of the task, whereas less successful participants disagreed with this statement. Only 8% (3/40) of the participants showed a clear discrepancy in the sense that they did not recognize their failure. Overall, 18% (7/40) of the participants did not recognize incomplete or irregular task fulfillment.

The answers were generally plausible and, in part, even elaborated.

Added Value of Our Study

This study is one of the few investigations of the interaction of people with dementia with an interactive smartwatch. Particularly noteworthy is the detailed clinical characterization of the study participants and the exact description of the technical features of the smartwatch interventions. The publication of the results of the survey about the wishes and needs of people with proven memory impairment regarding a smartwatch alone is valuable, even if the sample size of 40 individuals is not particularly large. Several specific use cases were mentioned, and new ideas for wishful smartwatch features were suggested.

Although there are numerous recommendations regarding the design of user interfaces of new technologies, for example, user interface design, these have rarely been tested in a scientific setting in practice on older adults or people with memory impairment. This target group is extremely heterogeneous, and
it must be assumed that general recommendations are of limited value [20]. Therefore, it was not clear whether a smartwatch would even be able to reach the threshold required to gain the attention and task understanding of these people.

Most studies of smartwatches for patients with dementia have no interactive claim but, instead, use passive monitoring [62]. We could only find 3 studies that used an interactive smartwatch for people with dementia to support activities of daily life [8,63,64]. König et al [8] designed a system in which a smartwatch was embedded into an infrastructure with tablets and a web platform to evaluate usability in a 3-month trial. The authors did not observe significant changes in the quantitative measurements. They did not assess single interactions. Thorpe et al [63] tested smartwatches capable of scheduling, navigation, communication, and orientation using an off-the-shelf wearable device in 5 participants. Personalization and familiarity appeared to be key drivers of smartwatch adoption. The task completion rates varied greatly among the participants, as in our study. McCarron et al [64] tested a face recognition application for smartwatches paired with smartphones to improve the quality of social interactions and quality of life among people with dementia. The authors had no trouble with study implementation and reported no problems with feasibility. Furthermore, they found no overall impact on the quality of life of the 48 participants when they used the smartwatch for 6 months.

Compared to other studies, our study has a sound sample size [12,15] and represents great added value, as it expands the small body of scientific literature in this specific area.

At a metalevel, this study is important because it highlights the practical implications of user-centered design in the development of novel technologies together with people with cognitive impairment, for example with giving precise recommendations for usability questionnaires for people with dementia, (Multimedia Appendix 6).

Limitations
Specific reasons for discomfort or concerns could not be captured using the questionnaire. In addition, contradictory information regarding the preferred intrusiveness of smartwatch notifications could not be resolved in this study. Some other questions remain open. What authority do we ascribe to technical devices? In this study, 1 participant (1/40, 3%) commented that they would not willingly drink the water that was provided even though the smartwatch prompted them to drink it. Another user stated that they were not thirsty. Such context-related information cannot be measured by sensors. If we want users to actively interact with a device, then questions about the immersive character arise. How should users know when they should interact with the smartwatch and when they should interact with their environment? The smartwatch specifications that we applied in our study seem to match the needs of users. Minor adjustments should focus on optimizing display latencies and wristbands. Future studies should implement more functions and test those functions in practice with more participants in real-world scenarios in the long term.

Conclusions
To determine whether user-centered design featuring people with cognitive impairment in the development of digital assistive devices is worthwhile, one needs to measure the utility of the developed technology. This can be accomplished by evaluating the usability, usefulness, and success metrics. All 3 approaches provide unique challenges because the available measuring tools may be inapplicable for the technology at hand or for the group of potential users. On the basis of the results of our study, we derived specific recommendations for questionnaires for people with dementia. Measures of internal consistency should not be overrated.

To improve technology adoption, the concept of “technical dyads” might be useful: each user is assigned to a person who is willing and capable of adjusting and maintaining devices for users who are technically inexperienced. This would be consistent with the needs of our participants, who expressed the wish for a permanent contact person for service. This wish also confirms the results that have already been found by others [10].

Our study is one of few studies that examined an interactive smartwatch for people with dementia. Although there are some methodological challenges for such studies, we and others have shown that both the inclusion of future users and the use of smartwatches by people with dementia are possible.

Acknowledgments
This study was funded by grants from the European Union (Europäischer Fonds für regionale Entwicklung [European Regional Development Fund]; TBI-V-1-100-VBW-035 and TBI-V-1-101-VBW-035). The funder played no role in study design, data collection, analysis and interpretation of data, or the writing of this paper. All authors declare no financial or non-financial competing interests. The authors have commissioned American Journal Experts, Durham, North Carolina, USA, to revise the language of the paper. The authors thank all participants for their participation and commitment.

Data Availability
The data sets used and analyzed during this study are available from the corresponding author upon reasonable request.

Authors’ Contributions
DG and ST conceptualized and designed the study with input from SK and GB. GB built the smartwatch prototype with feedback from DG and SK. DG, ER, and SK performed the study. DG analyzed the data. DG drafted the manuscript, which was edited and critically reviewed by ER, SK, GB, AGT, IK, and ST. ER contributed to literature review. AGT focused on language editing.
and coherence of content. SK and IK focused on shortening the manuscript and enhancing comprehensibility. All authors read and approved the final manuscript.

Conflicts of Interest

ST is a member of the advisory boards for Lilly, Eisai, and Biogen. He is also a member of the Independent Data Safety and Monitoring Board of the study ENVISION (Biogen). All the other authors have no conflicts to declare.

Multimedia Appendix 1
Sensor-Based Individualized Activity Management System for People With Dementia observational field study.

Multimedia Appendix 2
In-person testing of intervention tasks.

Multimedia Appendix 3
Demographics and test results.

Multimedia Appendix 4
Postexperiment questionnaire used in the intervention study.

Multimedia Appendix 5
Descriptive results of the applied questionnaire.

Multimedia Appendix 6
Recommendations for usability questionnaires for people with dementia.

References


Abbreviations

AD: Alzheimer disease
AT: assistive technology
JASP: Jeffrey’s Amazing Statistics Program
MCI: mild cognitive impairment
MMSE: Mini Mental Status Examination
SAMi: Sensor-Based Individualized Activity Management System for People With Dementia
SUS: System Usability Scale
TMT-A: Trail Making Test A
TMT-B: Trail Making Test B
Examining Associations Between Smartphone Use and Clinical Severity in Frontotemporal Dementia: Proof-of-Concept Study

Emily W Paolillo, PhD; Kaitlin B Casalotto, PhD; Annie L Clark, MS; Jack C Taylor, MAS; Hilary W Heuer, PhD; Amy B Wise, BS; Sreya Dhanam, BS; Mark Sanderson-Cimino, PhD; Rowan Saloner, PhD; Joel H Kramer, PsyD; John Kornak, PhD; Walter Kremers, PhD; Leah Forsberg, PhD; Brian Appleby, MD; Ece Bayram, MD, PhD; Andrea Bozoki, MD; Danielle Bruhaber, BS; R Ryan Darby, MD; Gregory S Day, MSc, MD; Bradford C Dickerson, MD; Kimiko Domoto-Reilly, MS, MD; Fanny Elahi, PhD; Julie A Fields, PhD; Nupur Ghoshal, MD, PhD; Neill Graff-Radford, MD; Matthew G H Hall, MS; Lawrence S Honig, MD; Edward D Huey, MD; Maria I Lapid, MD; Irene Litvan, MD; Ian R Mackenzie, MD; Joseph C Masdeu, MD, PhD; Mario F Mendez, MD, PhD; Carly Mester, BA; Toji Miyagawa, MD, PhD; Georges Naasan, MD, Belen Pascual, PhD; Peter Pressman, MD; Eliana Marisa Ramos, MD, PhD; Katherine P Rankin, PhD; Jessica Rexach, MD, PhD; Julio C Rojas, MD, PhD; Lawren VandeVrede, MD, PhD; Bonnie Wong, MD; Zbigniew K Wszolek, MD; Bradley F Boehle, MD; Howard J Rosen, MD; Adam L Boxer, MD, PhD; Adam M Staffaroni, PhD; ALLFTD Consortium

1Department of Neurology, Memory and Aging Center, Weill Institute for Neurosciences, University of California, San Francisco, San Francisco, CA, United States
2Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, CA, United States
3Department of Quantitative Health Sciences, Division of Clinical Trials and Biostatistics, Mayo Clinic, Rochester, MN, United States
4Department of Neurology, Mayo Clinic, Rochester, MN, United States
5Department of Neurology, Case Western Reserve University, Cleveland, OH, United States
6Department of Neurosciences, University of California, San Diego, La Jolla, CA, United States
7Department of Neurology, University of North Carolina, Chapel Hill, NC, United States
8Department of Neurology, Vanderbilt University, Nashville, TN, United States
9Department of Neurology, Mayo Clinic, Jacksonville, FL, United States
10Department of Neurology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, United States
11Department of Neurology, University of Washington, Seattle, WA, United States
12Department of Neurology, The Deane Center for Wellness and Cognitive Health, Icahn School of Medicine at Mount Sinai, New York, NY, United States
13Department of Psychiatry and Psychology, Mayo Clinic, Rochester, MN, United States
14Department of Neurology, Knight Alzheimer's Disease Research Center, Washington University, St. Louis, MO, United States
15Department of Neurology, Columbia University, New York, NY, United States
16Department of Psychiatry, Brown University, Providence, RI, United States
17Department of Pathology, University of British Columbia, Vancouver, BC, Canada
18Stanley H. Appel Department of Neurology, Nantz National Alzheimer Center, Houston Methodist Research Institute, Weill Cornell Medicine, Houston, TX, United States
19Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, CA, United States
20Department of Neurology, Icahn School of Medicine at Mount Sinai, New York, NY, United States
21Department of Neurology, University of Colorado, Aurora, CO, United States
22Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, Boston, MA, United States
23 see Acknowledgements, San Francisco, CA, United States

Corresponding Author:
Emily W Paolillo, PhD
Department of Neurology, Memory and Aging Center, Weill Institute for Neurosciences, University of California, San Francisco
675 Nelson Rising Lane, Suite 19094158
San Francisco, CA, 94158
United States
**Abstract**

**Background:** Frontotemporal lobar degeneration (FTLD) is a leading cause of dementia in individuals aged <65 years. Several challenges to conducting in-person evaluations in FTLD illustrate an urgent need to develop remote, accessible, and low-burden assessment techniques. Studies of unobtrusive monitoring of at-home computer use in older adults with mild cognitive impairment show that declining function is reflected in reduced computer use; however, associations with smartphone use are unknown.

**Objective:** This study aims to characterize daily trajectories in smartphone battery use, a proxy for smartphone use, and examine relationships with clinical indicators of severity in FTLD.

**Methods:** Participants were 231 adults (mean age 52.5, SD 14.9 years; n=94, 40.7% men; n=223, 96.5% non-Hispanic White) enrolled in the Advancing Research and Treatment of Frontotemporal Lobar Degeneration (ARTFL study) and Longitudinal Evaluation of Familial Frontotemporal Dementia Subjects (LEFFTDS study) Longitudinal Frontotemporal Lobar Degeneration (ALLFTD) Mobile App study, including 49 (21.2%) with mild neurobehavioral changes and no functional impairment (ie, prodromal FTLD), 43 (18.6%) with neurobehavioral changes and functional impairment (ie, symptomatic FTLD), and 139 (60.2%) clinically normal adults, of whom 55 (39.6%) harbored heterozygous pathogenic or likely pathogenic variants in an autosomal dominant FTLD gene. Participants completed the Clinical Dementia Rating plus National Alzheimer’s Coordinating Center Frontotemporal Lobar Degeneration Behavior and Language Domains (CDR+NACC FTLD) scale, a neuropsychological battery; the Neuropsychiatric Inventory; and brain magnetic resonance imaging. The ALLFTD Mobile App was installed on participants’ smartphones for remote, passive, and continuous monitoring of smartphone use. Battery percentage was collected every 15 minutes over an average of 28 (SD 4.2; range 14-30) days. To determine whether temporal patterns of battery percentage varied as a function of disease severity, linear mixed effects models examined linear, quadratic, and cubic effects of the time of day and their interactions with each measure of disease severity on battery percentage. Models covaried for age, sex, smartphone type, and estimated smartphone age.

**Results:** The CDR+NACC FTLD global score interacted with time on battery percentage such that participants with prodromal or symptomatic FTLD demonstrated less change in battery percentage throughout the day (a proxy for less smartphone use) than clinically normal participants (P<.001 in both cases). Additional models showed that worse performance in all cognitive domains assessed (ie, executive functioning, memory, language, and visuospatial skills), more neuropsychiatric symptoms, and smaller brain volumes also associated with less battery use throughout the day (P<.001 in all cases).

**Conclusions:** These findings support a proof of concept that passively collected data about smartphone use behaviors associate with clinical impairment in FTLD. This work underscores the need for future studies to develop and validate passive digital markers sensitive to longitudinal clinical decline across neurodegenerative diseases, with potential to enhance real-world monitoring of neurobehavioral change.

*(JMIR Aging 2024;7:e52831)* doi:10.2196/52831

**KEYWORDS**
digital; technology; remote; monitoring; cognition; neuropsychology; cognitive impairment; neurodegenerative; screening; clinical trials; mobile phone

**Introduction**

**Background**

Frontotemporal lobar degeneration (FTLD) is a common cause of dementia in individuals aged <65 years [1,2]. FTLD encompasses a group of neuropathologically distinct diseases that result in an overlapping set of dementia syndromes with heterogeneous symptoms, including those defined by primary behavior, language, or sensorimotor changes [3,4]. The timely detection of neurodegenerative diseases such as FTLD is a core public health strategy to reduce the individual, caregiver, and socioeconomic burden of dementia [5-7]. As we enter the era of disease-modifying treatments for neurodegenerative diseases, early detection is critical to identify those eligible for clinical trial participation and early treatment to slow or stop disease progression [8-10].

However, current assessment practices for detecting neurobehavioral changes associated with neurodegenerative disease are limited. In-person neuropsychological and neurological evaluations are the gold standard for determining the presence of cognitive impairment and identifying clinical phenotypes suggestive of an underlying neurodegenerative process; unfortunately, their high costs and restricted availability via specialty dementia clinics and research centers limit access for those with fewer financial resources and lower health literacy as well as those who reside in more remote geographic locations. In addition, evaluating a person at a single appointment provides only a snapshot of neurobehavioral functioning, which does not account for the dynamic nature of human behavior that fluctuates diurnally and is influenced by other dynamic factors (eg, sleep, fatigue, mood, and medications), limiting sensitivity for detecting early subtle declines [11,12]. Traditional neuropsychological assessment also lacks ecological validity
because interpretations of functioning are based on task performance in a tightly controlled testing environment, which seldom reflects a patient’s typical daily experience.

Remote monitoring of health status and behavior through the use of digital health tools is a promising solution to overcome the numerous limitations of in-person assessment and has been identified as a priority by several leading health organizations, including the US Food and Drug Administration [13], the US Department of Health and Human Services [14], and the National Institutes of Health [15,16]. Passive digital monitoring in particular (ie, monitoring behavior passively and unobtrusively through remote sensors) represents a low-burden and highly scalable method for improved detection and monitoring of real-world neurobehavioral change in neurodegenerative disease. Naturalistic behavioral data collected via in-home remote sensors have shown sensitivity to clinical severity in Alzheimer disease [17-25]; for example, older adults with mild cognitive impairment exhibit significant declines in the number of days with computer use and daily time spent on the computer per day compared to those without cognitive impairment [18]. As an extension of this work, we aim to examine overall daily smartphone use and its association with clinical severity in FTLD. We focused on FTLD as a specific use case to study the construct of passively collected smartphone data in the context of a neurodegenerative disease that manifests with well-characterized neurobehavioral changes.

Objectives
Thus, the aims of this study were to (1) examine passively collected battery percentage trajectories as a proxy for smartphone use throughout the day and (2) test associations between daily battery percentage trajectories and measures of cognitive and functional impairment and neurodegeneration in FTLD. Time-stamped battery percentage data can be easily accessed through public application programming interfaces (APIs) for both iOS and Android devices and have previously been associated with smartphone use [26-28]. Although smartphone screen time or app use time may be a more face-valid measure of smartphone use, access to these data has historically been restricted on iOS devices. This has been a major barrier to accessibility in passive monitoring research because nearly 30% of smartphone users worldwide have iOS devices [29]. Thus, it is worthwhile to examine battery percentage as a more accessible proxy for overall smartphone use. Consistent with prior research on computer use in older adults with cognitive impairment, we hypothesized that individuals with greater FTLD overall disease severity (ie, more severe functional impairment, worse cognitive performance, greater neuropsychiatric symptoms, and more brain atrophy) would demonstrate lower levels of daily smartphone use.

Methods
Participants
Participants were enrolled in the ARTFL (Advancing Research and Treatment of Frontotemporal Lobar Degeneration) study and LEFFTDS (Longitudinal Evaluation of Familial Frontotemporal Dementia Subjects) Longitudinal FrontoTemporal Lobar Degeneration (ALLFTD) Mobile App study through the multisite ALLFTD (NCT04363684) study and University of California San Francisco studies of FTLD (AG038791, AG062422, and AG019724), as described previously [30]. The participants were those who had a referring diagnosis of an FTLD clinical syndrome or those who were members of a family with a strong family history of an FTLD syndrome. Additional inclusion criteria were as follows: (1) aged ≥18 years, (2) access to a smartphone, and (3) English reported as the primary language. Participants were asked to use their own smartphones. Recruitment primarily targeted those with Clinical Dementia Rating Dementia Staging Instrument plus National Alzheimer’s Coordinating Center Frontotemporal Lobar Degeneration Behavior and Language Domains (CDR+NACC FTLD) global scores of <2, but participants who were more severely impaired were not excluded. Data for this study were collected from August 2020 to April 2023. During this period, 257 participants were enrolled and logged into the ALLFTD Mobile App on their personal smartphones. Participants were only included in this secondary analysis of the ALLFTD Mobile App study if they had at least 14 continuous days of passive smartphone monitoring data, consistent with prior digital phenotyping studies attempting to capture typical daily behavior [31]. Thus, of the initial 257 participants, 231 (89.9%) were included in the final sample after 26 (10.1%) participants were excluded because they first logged in <14 days before the date on which these data were pulled in April 2023. Of these 231 participants, 92 (39.8%) were classified as having neurobehavioral symptoms at the prodromal stage (ie, no functional impairment) or fully symptomatic (ie, with functional impairment) level of severity that are consistent with an FTLD-related clinical phenotype per conference consensus with neurologists and neuropsychologists following published criteria [32-35]. Participants who were symptomatic had either sporadic FTLD or a confirmed pathogenic or likely pathogenic variant in an autosomal dominant FTLD gene (ie, a pathogenic expansion in the chromosome 9 open reading frame 72 [C9orf72] gene or a known pathogenic or likely pathogenic variant in the progranulin (GRN) or microtubule-associated protein tau [MAPT] genes; conducted as described previously [36]). The remaining participants (139/231, 60.2%) were asymptomatic clinically normal family members of the prodromal or symptomatic individuals who (1) carried a pathogenic or likely pathogenic FTLD gene variant (55/139, 39.6%), (2) tested negative for known pathogenic or likely pathogenic FTLD variants (50/139, 36%), or (3) did not yet have results available from genetic testing (34/139, 24.5%).

Ethical Considerations
The study was approved by a centralized single institutional review board at Johns Hopkins Medicine (IRB # 20-29891), and all participants provided written informed consent.

Measures
Passively Monitored Smartphone Battery Percentage
We used the first 30 days of participants’ smartphone data for this study with the goal of understanding whether approximately 1 month of smartphone monitoring could reflect baseline neurobehavioral status without capturing longitudinal disease-related decline [18,37]. Participants downloaded the
ALLFTD Mobile App onto their personal smartphones. The app is designed to deliver both active mobile cognitive assessments and passively collect smartphone use data [30], including battery percentage. The ALLFTD Mobile App was programmed to collect battery percentage every 15 minutes. Due to some variability around this timing in the actual data collected (ie, some missing data points and some data collected over shorter intervals), data were aggregated to reflect the average battery percentage per hour of each study day per participant. This resulted in a comparable number of data points per day across participants. The ALLFTD Mobile App also recorded information about participants’ smartphone model, which was used to estimate the age of the smartphone (ie, calculated on the basis of the smartphone model release date and the first date of participation in this study).

Functional, Cognitive, and Neuropsychiatric Assessment
All participants underwent comprehensive functional and cognitive assessment at a parent study visit at the beginning of their smartphone monitoring study period. Informant and participant interviews were conducted to characterize the level of cognitive and everyday functioning impairment using the CDR+NACC FTLD scale [38], which is a validated, modified version of the CDR [39] that has higher sensitivity to functional impairment in FTLD. CDR+NACC FTLD global scores [40] were used to categorize participants into disease severity groups: 0=unimpaired, 0.5=prodromal, and ≥1=symptomatic. Domain-specific cognitive functioning was assessed via a comprehensive battery of well-validated neuropsychological tests. The previously published Uniform Data Set (Version 3) Executive Function composite score was used as our measure of executive functioning, comprising Trail Making Test A and B, phonemic fluency (generating words beginning with F and L), number span backwards, and category fluency (animals and vegetables) [41,42]. Sample-based z scores were calculated for indices of memory, including immediate and delayed free recall on the California Verbal Learning Test-3 Brief Form [43], as well as Benson Complex Figure Delayed Recall [44,45]. A composite memory z score was created by taking the mean of the z scores across these memory tests. Language functioning was assessed via the Multilingual Naming Test [46]. Visuospatial functioning was assessed via the Benson Complex Figure Copy [45]. Informants also completed the Neuropsychiatric Inventory [47] to assess the presence and severity of neuropsychiatric symptoms in participants.

Neuroimaging
Of the 231 participants, a subset (n=189, 81.8%) completed neuroimaging. Participants were scanned on 3 Tesla magnetic resonance imaging (MRI) scanners. T1-weighted images were acquired as magnetization-prepared rapid gradient echo images using the following parameters: 240×256×256 matrix; approximately 170 slices; voxel size=1.05×1.05×1.25 mm³; and flip angle, echo time, and repetition time varied by vendor. A standard imaging protocol was used across all centers, and all images were reviewed for quality by a core group at the Mayo Clinic, Rochester, Minnesota, United States. Details of image acquisition, processing, and harmonization have been published elsewhere [48]. Total gray matter volume was used as the primary neuroimaging variable of interest. Total intracranial volumes were regressed out (using a simple linear regression with gray matter volume as outcome and total intracranial volume as the only predictor) before inclusion in analyses to account for interindividual volumetric differences in head size on gray matter volume.

Statistical Analyses
Differences in demographic and clinical characteristics across the CDR+NACC FTLD–defined disease severity groups were tested with 1-way ANOVA and chi-square tests for continuous variables and categorical variables, respectively. Raw battery percentage data were plotted against the time of day to inform statistical analysis. Linear mixed effects (LME) regression models were then used to model the linear, quadratic, and cubic effects of time (ie, hour of the day; 0=midnight; 23=11 PM) on battery percentage. Person-specific random intercepts and random effects of time (linear, quadratic, and cubic) were modeled. To determine whether daily patterns of battery percentage trajectories (ie, a proxy for smartphone use) varied as a function of FTLD disease severity, LME models examined linear, quadratic, and cubic effects of time and their interaction with each measure of disease severity separately (ie, CDR+NACC FTLD group, cognitive domain z scores, neuropsychiatric symptoms, and whole brain gray matter volumes). All LME models covaried for age, sex, smartphone type (iOS vs Android), and estimated smartphone age (calculated on the basis of the smartphone model release date and the first date of participation in this study). A post hoc sensitivity analysis was conducted in a subset of participants (162/231, 70.1%) whose age range was matched across the CDR+NACC FTLD groups. To understand whether subtle differences in neurobehavioral functioning related to daily smartphone battery use trajectories in an unimpaired sample, we conducted additional sensitivity analyses, which repeated all models in the subset of participants who were clinically normal (139/231, 60.2%). Regression estimates are reported as standardized betas, which represent the predicted change in the outcome as a function of each predictor in units of SDs. All analyses were performed using R (version 4.2.0; R Foundation for Statistical Computing). The lme4 package was used to conduct the LME regressions [49].

Results

Participant Characteristics
Table 1 shows demographic and clinical characteristics by disease severity group. Participants had a mean age of 52.5 (SD 14.9) years and a mean of 16 (SD 2.2) years of education. Of the 231 participants, 94 (40.7%) were men, and 223 (96.5%) identified as non-Hispanic White. Nearly three-fourths of the participants (171/231, 74%) were women (P<.001) and more likely to be women (P<.001) than those with prodromal or symptomatic FTLD, consistent with the larger parent study samples (ALLFTD Mobile App study [30]; ALLFTD [40,50]). There were no other clear imbalances in demographics.
other demographic and clinical characteristics across the 3 groups. Overall, participants had a mean of 28.3 (SD 4.19; range 14-30) days of smartphone monitoring data. On average, participants’ smartphones were 2.8 (SD 1.53; range 0-7) years old.
Table 1. Participant characteristics by disease severity group (n=231).

<table>
<thead>
<tr>
<th>Demographics</th>
<th>A (clinically normal; n=139)</th>
<th>B (prodromal; n=49)</th>
<th>C (symptomatic; n=43)</th>
<th>P value</th>
<th>Pairwise comparisonsa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>46.3 (13.9)</td>
<td>59.7 (12.1)</td>
<td>64.3 (9.3)</td>
<td>&lt;.001</td>
<td>A&lt;B, C</td>
</tr>
<tr>
<td>Sex (male), n (%)</td>
<td>41 (29.5)</td>
<td>30 (61.2)</td>
<td>23 (53.5)</td>
<td>&lt;.001</td>
<td>A&lt;B, C</td>
</tr>
<tr>
<td>Education (years), mean (SD)</td>
<td>16.3 (2.1)</td>
<td>16.5 (2.4)</td>
<td>16.7 (2.4)</td>
<td>.61</td>
<td>N/Ab</td>
</tr>
<tr>
<td>Race and ethnicity (non-Hispanic White), n (%)</td>
<td>134 (96.4)</td>
<td>48 (98)</td>
<td>41 (95.3)</td>
<td>.79</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Study characteristics

| Total study days, mean (SD) | 28.3 (4.2) | 28.1 (4.3) | 28.6 (4.2) | .88 |
| Smartphone type, n (%) | .42 |
| iOS | 97 (69.8) | 37 (75.5) | 27 (62.8) |
| Android | 42 (30.2) | 12 (24.5) | 16 (37.2) |
| Estimated smartphone age (years), mean (SD) | 2.7 (1.5) | 3.0 (1.5) | 2.9 (1.6) | .30 |

Genetic status

| Genetic testing results, n (%) | .56 |
| Not available | 34 (24) | 13 (26.5) | 13 (30.2) |
| Available | 105 (75.5) | 36 (73.5) | 30 (69.8) |
| Mutation carrier | |
| C9orf72c | 55 (52.4) | 15 (41.7) | 7 (23.3) |
| GRNd | 29 (52.7) | 8 (53.3) | 3 (42.9) |
| MAPTe | 7 (12.7) | 1 (6.7) | 0 (0) |
| Otherf | 16 (29.1) | 6 (40) | 3 (42.9) |

Clinical phenotype

| Mild cognitive impairmentifa | N/A | 39 (79.6) | N/A |
| bvFTDfb | N/A | N/A | 25 (58.1) |
| svPPAfa | N/A | N/A | 3 (7) |
| nfvPPAf | N/A | N/A | 3 (7) |
| lvPPAk | N/A | N/A | 1 (2.3) |
| PSP-RSl | N/A | 3 (6.1) | 4 (9.3) |
| CBSm | N/A | 2 (4.1) | 2 (4.7) |
| Othern | N/A | 5 (10.2) | 2 (4.7) |

aPairwise comparisons were evaluated with the Tukey honestly significant difference test.
bN/A: not applicable.
cC9orf72: chromosome 9 open reading frame 72.
dGRN: progranulin.
eMAPT: microtubule-associated protein tau.
fIdentified pathogenic or likely pathogenic variants in genes less commonly identified as genetic causes of frontotemporal lobar degeneration (FTLD; ie, other than C9orf72, GRN, or MAPT). The specific genetic variant is not provided to protect participant anonymity.
fIncludes behavior-, cognitive-, and language-predominant mild cognitive impairment syndromes.
hbvFTD: behavioral variant frontotemporal dementia.
iSVPPA: semantic variant primary progressive aphasia.
jnfvPPA: nonfluent variant primary progressive aphasia.
kLvPPA: logopenic variant primary progressive aphasia.

https://aging.jmir.org/2024/1/e52831

JMIR Aging 2024 | vol. 7 | e52831 | p.704

XSL-FO

RenderX
Daily Smartphone Battery Percentage
Visualization of the raw battery percentage data by the time of day (Figure 1) shows a nonlinear trajectory such that, on average, battery percentage increased from midnight to approximately 6 AM, then decreased until about 7 PM, and then increased again through 11 PM. These temporal patterns presumably represent typical patterns of charging and charge use of the smartphone throughout the day. Multimedia Appendix 1 presents raw battery percentage data by disease severity group. The shape of these raw data motivated consideration of a cubic model. Thus, we first tested the fit of the LME regression modeling the linear, quadratic, and cubic effects of the time of day on battery percentage, covarying for age, sex, smartphone type, and estimated smartphone age. The cubic model’s conditional pseudo-$R^2$ (ie, the proportion of variance explained by both fixed and random factors) was 0.37. The likelihood ratio tests indicated that the full cubic model had statistically significantly better fit than LME regressions modeling only the linear ($\chi^2=4283.6; P<.001$) and quadratic ($\chi^2=4118.8; P<.001$) effects of time.

LME regression indicated that the interactions between disease severity group and the linear, quadratic, and cubic effects of the time of day were associated with battery percentage (Table 2). Visualization of model results suggests that participants with prodromal FTLD and those with symptomatic FTLD had flatter battery curves throughout the day (ie, shallower decreases from maximum to minimum battery percentage as well as a higher minimum battery percentage; a proxy for less smartphone use) than clinically normal participants on average (Figure 2). Examination of pairwise disease severity group contrasts showed that participants with symptomatic FTLD also had significantly less battery use than participants with prodromal FTLD ($P=.003$ or $P<.001$ in all cases).

Figure 1. Visualization of raw battery percentage data for all participants binned by time of day (0=midnight; 23=11 PM).
Table 2. Linear mixed effects regression results showing significant relationships between disease severity groups and battery percentage trajectories throughout the day.

<table>
<thead>
<tr>
<th></th>
<th>$\beta$ (95% CI)</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline age</td>
<td>.06 (.00 to .13)</td>
<td>.047</td>
</tr>
<tr>
<td>Sex (reference: female)</td>
<td>−.03 (−.14 to .08)</td>
<td>.62</td>
</tr>
<tr>
<td>Smartphone type (reference: Android)</td>
<td>.01 (−.11 to .12)</td>
<td>.89</td>
</tr>
<tr>
<td>Estimated smartphone age</td>
<td>−.04 (−.08 to −.01)</td>
<td>.03</td>
</tr>
<tr>
<td>Time of day (linear)</td>
<td>1.74 (1.66 to 1.83)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (quadratic)</td>
<td>−5.27 (−5.45 to −5.09)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (cubic)</td>
<td>3.28 (3.17 to 3.40)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Prodromal (reference: normal)</td>
<td>.24 (.10 to .39)</td>
<td>.001</td>
</tr>
<tr>
<td>Symptomatic (reference: normal)</td>
<td>.27 (.11 to .43)</td>
<td>.001</td>
</tr>
<tr>
<td>Time of day (linear) × prodromal</td>
<td>−.23 (−.40 to −.07)</td>
<td>.006</td>
</tr>
<tr>
<td>Time of day (linear) × symptomatic</td>
<td>−.56 (−.73 to −.38)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (quadratic) × prodromal</td>
<td>.88 (.53 to 1.24)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (quadratic) × symptomatic</td>
<td>1.65 (1.29 to 2.01)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (cubic) × prodromal</td>
<td>−.58 (−.81 to −.36)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (cubic) × symptomatic</td>
<td>−.99 (−1.22 to −.76)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Figure 2. Participants with prodromal frontotemporal lobar degeneration (FTLD) and those with symptomatic FTLD had flatter battery curves throughout the day (ie, a proxy for less smartphone use) than clinically normal participants on average. Error bands represent pointwise 95% CIs.

Similar patterns emerged when examining all other indices of clinical severity. Each cognitive domain statistically significantly moderated the relationship between the time of day and battery percentage such that participants with worse cognitive functioning had flatter battery curves throughout the day, suggesting less smartphone use (Table 3 [executive functioning, memory, language, and visuospatial skills]; Figures 3A-3D). Neuropsychiatric symptom severity also moderated the relationship between the time of day and battery percentage such that participants with higher neuropsychiatric symptom ratings had flatter battery curves throughout the day, suggesting less smartphone use (Table 3 [neuropsychiatric symptoms]; Figure 3E). Examination of each Neuropsychiatric Inventory item (yes or no) in separate LME models suggested that participants with agitation, depression, apathy, disinhibition, irritability, motor disturbance, nighttime behaviors, and changes in appetite had less smartphone use (Table 4). Delusions, hallucinations, anxiety, and elation did not statistically
significantly relate to battery use trajectories throughout the day (Table 4). Finally, total gray matter volume also moderated the relationship between the time of day and battery percentage such that participants with smaller gray matter volumes had flatter battery curves throughout the day, suggesting less smartphone use (Table 3 [gray matter volume]; Figure 3F). Of all indices of clinical severity presented in Table 3, executive functioning and total gray matter volume appeared to have the largest effect sizes on smartphone battery trajectories.

Given the age difference across disease severity groups, we repeated the first LME model examining battery percentage trajectories by CDR+NACC FTLD group after restricting the age range of the clinically normal group to be identical to that of the group with prodromal FTLD and the group with symptomatic FTLD (participants aged 44-81 years in all groups; clinically normal: 70/139, 50.4%). The interactions between disease severity group and the linear, quadratic, and cubic effects of the time of day on battery percentage are fairly consistent, showing that the participants who were symptomatic had lower battery use than clinically normal participants (interaction with linear time: $\beta = -0.23, 95\% CI \text{-.40 to -.06}; P = .009$; interaction with quadratic time: $\beta = 0.65, 95\% CI \text{.23 to 1.07}; P = .003$; interaction with cubic time: $\beta = -0.39, 95\% CI \text{-66 to -12}; P = .004$). However, the difference between the prodromal and clinically normal participants no longer reached statistical significance (interaction with linear time: $\beta = -0.05, 95\% CI \text{-.13 to .22}; P = .586$; interaction with quadratic time: $\beta = -0.07, 95\% CI \text{-50 to .36}; P = .76$; interaction with cubic time: $\beta = 0.05, 95\% CI \text{-22 to .33}; P = .70$).

Table 3. Results of separate linear mixed effects regression models showing significant relationships between battery percentage trajectories throughout the day and executive functioning, memory, language, visuospatial skills, neuropsychiatric symptoms, and whole brain gray matter volume (lower order terms and covariates are not displayed).

<table>
<thead>
<tr>
<th></th>
<th>$\beta$ (95% CI)</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Executive functioning</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of day (linear)×UDS3-EF$^a$ composite score</td>
<td>.24 (.17 to .31)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (quadratic)×UDS3-EF composite score</td>
<td>-.65 (-.80 to -.50)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (cubic)×UDS3-EF composite score</td>
<td>.38 (.28 to .47)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Memory</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of day (linear)×memory $z$ score</td>
<td>.21 (.13 to .28)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (quadratic)×memory $z$ score</td>
<td>-.58 (-.74 to -.42)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (cubic)×memory $z$ score</td>
<td>.34 (.24 to .44)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of day (linear)×MINT$^b$ $z$ score</td>
<td>.07 (.00 to .15)</td>
<td>.05</td>
</tr>
<tr>
<td>Time of day (quadratic)×MINT $z$ score</td>
<td>-.24 (-.40 to -.08)</td>
<td>.002</td>
</tr>
<tr>
<td>Time of day (cubic)×MINT $z$ score</td>
<td>.15 (.05 to .25)</td>
<td>.004</td>
</tr>
<tr>
<td><strong>Visuospatial skills</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of day (linear)×Benson Complex Figure Copy $z$ score</td>
<td>.10 (.04 to .16)</td>
<td>.001</td>
</tr>
<tr>
<td>Time of day (quadratic)×Benson Complex Figure Copy $z$ score</td>
<td>-.28 (-.38 to -.19)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (cubic)×Benson Complex Figure Copy $z$ score</td>
<td>.19 (.13 to .25)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Neuropsychiatric symptoms</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of day (linear)×NPI total score</td>
<td>-.06 (-.13 to .01)</td>
<td>.08</td>
</tr>
<tr>
<td>Time of day (quadratic)×NPI total score</td>
<td>.25 (.10 to .39)</td>
<td>.001</td>
</tr>
<tr>
<td>Time of day (cubic)×NPI total score</td>
<td>-.15 (-.25 to -.05)</td>
<td>.002</td>
</tr>
<tr>
<td><strong>Gray matter volume (n=189)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time of day (linear)×gray matter volume</td>
<td>.22 (.14 to .29)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (quadratic)×gray matter volume</td>
<td>-.73 (-.88 to -.57)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Time of day (cubic)×gray matter volume</td>
<td>.45 (.35 to .55)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

$^a$UDS3-EF: Uniform Data Set (Version 3) Executive Function.

$^b$MINT: Multilingual Naming Test.

$^c$NPI: Neuropsychiatric Inventory.
Figure 3. Daily battery percentage trajectories were significantly moderated by (A) executive functioning, (B) memory, (C) language, (D) visuospatial skills, (E) neuropsychiatric symptoms, and (F) total gray matter volumes. Participants with worse neurobehavioral outcomes had smaller daily decreases from peak to minimum battery percentage on average, suggesting less smartphone use throughout the day. GMV: gray matter volume; MINT: Multilingual Naming Test; NPI: Neuropsychiatric Inventory; TIV: total intracranial volume; UDS3-EF: Uniform Data Set (Version 3) Executive Function.

Table 4. Results of separate linear mixed effects regression models examining relationships between battery percentage trajectories throughout the day and each neuropsychiatric symptom (NPS) captured on the Neuropsychiatric Inventory. Models covaried for age, sex, smartphone type, and estimated smartphone age.

<table>
<thead>
<tr>
<th>Predictors modeling interactions with the time of day</th>
<th>NPS×time of day (linear), β (SE)</th>
<th>NPS×time of day (quadratic), β (SE)</th>
<th>NPS×time of day (cubic), β (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delusions</td>
<td>−.12 (.25)</td>
<td>−.25 (.53)</td>
<td>.43 (.34)</td>
</tr>
<tr>
<td>Hallucinations</td>
<td>−.61 (.40)</td>
<td>1.62 (.87)</td>
<td>−.78 (.55)</td>
</tr>
<tr>
<td>Agitation</td>
<td>−.36 (.10)</td>
<td>1.06 (.20)</td>
<td>−.65 (.13)</td>
</tr>
<tr>
<td>Depression</td>
<td>−.34 (.08)</td>
<td>.92 (.17)</td>
<td>−.64 (.11)</td>
</tr>
<tr>
<td>Anxiety</td>
<td>.05 (.08)</td>
<td>−.20 (.17)</td>
<td>.15 (.11)</td>
</tr>
<tr>
<td>Elation</td>
<td>.12 (.13)</td>
<td>−.38 (.26)</td>
<td>.29 (.17)</td>
</tr>
<tr>
<td>Apathy</td>
<td>−.43 (.08)</td>
<td>1.32 (.17)</td>
<td>−.81 (.11)</td>
</tr>
<tr>
<td>disinhibition</td>
<td>−.21 (.09)</td>
<td>1.14 (.18)</td>
<td>−.86 (.12)</td>
</tr>
<tr>
<td>Irritability</td>
<td>−.20 (.07)</td>
<td>.78 (.16)</td>
<td>−.49 (.10)</td>
</tr>
<tr>
<td>Motor disturbance</td>
<td>−.44 (.11)</td>
<td>1.18 (.24)</td>
<td>−.69 (.15)</td>
</tr>
<tr>
<td>Nighttime behaviors</td>
<td>−.41 (.10)</td>
<td>1.20 (.22)</td>
<td>−.74 (.14)</td>
</tr>
<tr>
<td>Changes in appetite</td>
<td>−.16 (.09)</td>
<td>.86 (.20)</td>
<td>−.61 (.13)</td>
</tr>
</tbody>
</table>

aP values met the threshold for significance.

Sensitivity analyses conducted among the 139 clinically normal participants showed that the following neurobehavioral measures were associated with daily battery percentage trajectories: executive functioning (interaction with linear time: β=.21, 95% CI .12 to .30; P<.001; interaction with quadratic time: β=.57, 95% CI −.76 to −.38; P<.001; interaction with cubic time: β=.34, 95% CI .21 to .46; P<.001); memory (interaction with linear time: β=.33, 95% CI .24 to .42; P<.001; interaction with quadratic time: β=.83, 95% CI −1.02 to −.63; P<.001; interaction with cubic time: β=.47, 95% CI .35 to .60; P<.001), and total gray matter volume (interaction with linear time: β=.32, 95% CI .22 to .41; P<.001; interaction with quadratic time: β=.57, 95% CI .36 to .78; P<.001).
β=-.84, 95% CI −1.03 to −.64; P<.001; interaction with cubic time: β=.48, 95% CI .36 to .61; P<.001). Language, visuospatial functioning, and neuropsychiatric symptoms did not strongly associate with daily battery percentage trajectories in clinically normal participants. The directions of associations in these clinically normal participants were consistent with relationships described in the entire sample.

**Discussion**

**Principal Findings**

This study is the first to our knowledge to examine passively collected smartphone use data in a sample with neurodegenerative disease. The results highlight an accessible, low-burden, and scalable remote monitoring method that captured behaviors associated with cognitive, neuropsychiatric, and brain health outcomes in a sample of participants with FTLD. The findings support a proof of concept that this passive digital monitoring approach, in combination with other methods, warrants further evaluation as a potential tool to augment screening and monitoring neurobehavioral change in clinical populations. Consistent with our hypotheses, we found that daily trajectories of smartphone battery use (a proxy for overall smartphone use) were associated with gold standard measures of clinical severity in FTLD such that those with more severe levels of impairment had less smartphone use throughout the day. Relationships between battery percentage trajectories and executive functioning, memory, and gray matter volume also held in the subset of clinically normal participants, suggesting potential sensitivity to subclinical neurobehavioral differences.

**Comparison to Prior Work**

These findings are consistent with previous studies showing that older adults with cognitive impairment have greater declines in everyday technology use compared to cognitively unimpaired older adults [18,51,52]. Other studies have shown that older adults with cognitive impairment report more difficulties using technology, representing a potential barrier to technology use [53,54]. Notably, the observed associations between battery percentage trajectories and cognitive functioning were not specific to particular cognitive domains, suggesting that the metrics of overall smartphone use may reflect a global transdiagnostic marker of functioning rather than a phenotype-specific marker (e.g., executive functioning–predominant or language–predominant dysfunction). Thus, our findings may also not be specific to FTLD, and future work is needed to replicate findings in other populations with neurologic conditions. The use of a smartphone, like the use of a computer [18], is a cognitively complex task requiring the resources of many functions (e.g., attention, executive function, working memory, and fine motor skill). As such, smartphone use patterns may be a particularly sensitive marker of early and subtle neurobehavioral change; however, additional research examining longitudinal changes in smartphone use over time is needed to support this hypothesis.

While this is the first study to our knowledge to report on passively collected smartphone use data in the context of neurodegenerative disease, there is a growing body of literature examining other passive streams of smartphone data as potential markers of neurobehavioral function in older adults; for example, passively collected data from smartphone accelerometers, GPS location, and touchscreen typing have been associated with symptom and disease severity in Parkinson disease, multiple sclerosis, and amyotrophic lateral sclerosis [55-64]. Future work should incorporate multiple passive smartphone data types for more comprehensive digital phenotyping and potentially improved clinical relevance in monitoring neurodegenerative disease.

Regarding the more technological aspects of passive smartphone data collection, previous studies have also reported similar variability around the frequency and timing of data collected per person. These studies have identified a number of factors that influence the collection and transfer of smartphone data to secure cloud-based servers, including smartphone hardware, data permissions, app engagement, wireless service, capacity of local data storage, data transmission limits, and even sociodemographic factors [65-67]. This has also been reported in other devices as well, including wearables [68-70]. Thus, thorough data cleaning is necessary to ensure that enough data points are captured to accurately represent activity for a given time period, as has been described previously [71,72].

Visualizing the raw battery percentage data was important for understanding daily patterns and supported the utility of daily smartphone battery percentage trajectories as a proxy for smartphone use. Average patterns in battery percentage appeared to track with typical diurnal sleep-wake rhythms: percentages increased up to morning (when mobile phones are likely charging) and decreased throughout the day (when participants were presumably awake and using their smartphones) until nighttime when percentages began to increase again. Careful examination of these raw data patterns, alongside measures of clinical severity, may support the development of specific metrics using battery percentage data that can be easily tracked over longer periods of time (e.g., total battery drainage per day). Future work in this field should also consider examining the frequency and timing of smartphone battery charging as a way to track routine daily use patterns that may be clinically relevant. Tracking these metrics over many months or years would allow for future studies to examine person-specific changes in battery use over time and test longitudinal associations with neurodegenerative disease progression.

**Strengths and Limitations**

The strengths of this study include a large, extensively characterized cohort with FTLD; the reporting of novel smartphone use data; and the use of passive digital monitoring techniques. However, we also acknowledge several limitations. First, there are certainly caveats to our approach using battery percentage as a proxy for smartphone use, including factors that are difficult to quantify and adjust for, such as the impact of hardware, software, and service connection on battery life [73]. In addition, the ALLFTD Mobile App does not record when a smartphone may have been plugged in for charging, potentially preventing periods of battery decline even when the smartphone may have been in use. Even so, the robust relationships observed with gold standard measures of functional impairment, cognition, neuropsychiatric symptoms, and brain volumes are...
encouraging. Second, it is likely that the mobile cognitive testing sessions administered by the ALLFTD Mobile App contributed to some smartphone battery use and that battery use may subsequently be affected by adherence to the mobile cognitive testing protocol. However, we have previously reported on adherence to mobile cognitive testing through this app [30], which showed that cognitive testing completion rates among asymptomatic participants, participants with prodromal FTLD, and participants with symptomatic FTLD were 71.4%, 78.4%, and 59%, respectively. These adherence rates do not match the stairstep effect of battery use reported in this study whereby cognitively normal participants exhibited the highest smartphone battery use, participants with prodromal FTLD demonstrated intermediate smartphone battery use, and participants with symptomatic FTLD showed the lowest smartphone battery use, suggesting that our results are not simply driven by adherence to the mobile cognitive testing protocol. Third, although we importantly controlled for the effects of age, sex, smartphone type, and estimated smartphone age in our statistical analyses, further replication is needed in samples whose disease severity groups are demographically matched. Fourth, it is possible that a 30-day monitoring period may not be enough time to most accurately capture routine smartphone use behavior. Future studies are needed to evaluate the psychometrics of passive smartphone use metrics across different periods of time to identify optimal lengths of follow-up. Fifth, our sample was also limited in demographic diversity because participants mostly identified as non-Hispanic White and were highly educated, reflective of the cohort in the ALLFTD study [10]. This is a crucial consideration when examining new tools that require access to technology because the implementation of such digital tools may inadvertently increase disparities among those with fewer resources to obtain technology that meets required software specifications. However, with the steadily increasing rates of smartphone ownership worldwide in addition to the implementation of government-funded programs to provide access to technology (eg, the Lifeline Program or “Obama Phone” [74]), there is growing consensus that smartphone monitoring could become universally accessible.

Conclusions

In sum, our novel results demonstrate the feasibility of continuous, unobtrusive smartphone use monitoring, while also showing that smartphone use relates to the severity of neurobehavioral impairment in a sample with FTLD. We highlight these results as proof of concept because we believe that they support future research examining whether specific smartphone use metrics are clinically relevant and may have utility for monitoring clinical disease progression in FTLD and other neurodegenerative diseases. With continued validation, such passive monitoring methodologies for real-time, real-world, and remote monitoring have the potential to improve the monitoring of clinically meaningful neurobehavioral changes in individuals at risk for dementia.

Acknowledgments

This work was supported by the National Institutes of Health (AG063911, AG077557, AG62677, AG045390, NS092089, AG032306, AG021886, AG016976, AG058233, AG038791, AG02350, AG019724, AG062422, NS050915, AG032289-11, AG077557, K23AG061253, K24AG045333, K23AG59888, K23AG064029, and K23AG073514). This work was also supported by the Association for Frontotemporal Degeneration, the Bluefield Project to Cure Frontotemporal Dementia (FTD), the Rainwater Charitable Foundation, AlzOut, the John Douglas French Alzheimer’s Foundation, and the Alzheimer’s Association (AARP-22-974065, principal investigator: EWP; AARP-23-1145318, principal investigator: RS). The authors thank the organizers of the the International Society to Advance Alzheimer’s Research and Treatment (ISTAART) Design and Data Analytics Professional Interest Area Workshop, Applying Novel Models for Longitudinal Outcomes, who provided consultation on the statistical analyses in this study.

The members of ALLFTD Consortium are Liana Apostolova, MD, Indiana University; Sami Barrameda, MD, PhD, University of Michigan; Ece Bayram, MD, PhD, UCSD; Hugo Botha, MD, Mayo Clinic, Rochester; Andrea Bozoki, MD, University of North Carolina; Yann Cobigo, PhD, UCSF; Gregg S Day MD, MSc, Mayo Clinic Jacksonville; Dennis Dickinson, MD, Mayo Clinic, Jacksonville; Kelley Faber, MS, Indiana University (NCRAD); Anne Fagan, PhD, Washington University, St. Louis; Jamie Fong, MS, UCSF; Tatiana Foroud, PhD, Indiana University (NCRAD); Douglas R. Galasko, MD, UCSD; Ralitza Gavrilova, MD, Mayo Clinic, Rochester; Tania Gendron, PhD, Mayo Clinic Jacksonville; Daniel Geschwind, MD, PhD, UCLA; Jill Goldman, MS, MPhil, Columbia University; Jonathan Graff-Radford, MD, Mayo Clinic, Rochester; Ian M. Grant, MD, Northwestern University; Murray Grossman, MD, EdD, University of Pennsylvania; Chadwick M. Hales, MD, PhD, Emory; Hilary W. Heuer, PhD, UCSF; Ying-Yuek (Robin) Hsiung, MD, University of British Columbia; Eric Huang, PhD, UCSF; David Irwin, MD, University of Pennsylvania; Noah Johnson, BA, Mayo Clinic, Rochester; David T. Jones, MD, Mayo Clinic, Rochester; Kajal Kantarci, MD, Mayo Clinic, Rochester; David Knopman, MD, Mayo Clinic, Rochester; Tyler Kolander, BA, Mayo Clinic, Rochester; Justin Kwan, MD, National Institutes of Health; Argentinia Lario Lago, PhD, UCSF; Shannon B. Lavigne, PhD, UT Health San Antonio; Suzeet Lee, MD, UCSF; Gabriel C. Léger, MD, UCSD; Peter Ljubenkov, MD, UCSF; Diane Lucente, MS, MGH; Masood Manochehr, BA, Columbia University; Lauren Massimo, PhD, University of Pennsylvania; Scott McGinnis, MD, MGH; Corey T. McMillan, PhD, University of Pennsylvania; Joie Molen, PhD, University of Colorado; Chiang Ouyouke, MBBS, MHS Johns Hopkins University; Alexander Panteleyt, MD, Johns Hopkins University; Henry Paulson, MD, PhD, University of Michigan; Leonard Petrucelli, PhD, Mayo Clinic Jacksonville; Rosa Rademakers, PhD, University of Antwerp; Vijay Ramanan MD, PhD, Mayo Clinic, Rochester; Meghanha Rao, BS, MPH Mayo Clinic, Rochester; Katya Rasovsky, PhD, University of Pennsylvania; Kristoffer W. Rhoads, PhD, University of Washington; Aaron Ritter, MD, Cleveland Clinic Las Vegas; Erik D. Roberson MD,
PhD, University of Alabama at Birmingham; Emily Rogalski, PhD, Northwestern University; Rodolfo Savica, MD, PhD, Mayo Clinic, Rochester; William Seeley, MD, UCSF; Allison Snyder, MD, National Institutes of Health; Anne C. Sullivan, PhD, UT Health San Antonio; Jeremy M. Syrjanen, MS, Mayo Clinic, Rochester; M. Carmela Tartaglia, MD, University of Toronto; Philip W. Tipton, MD, Mayo Clinic Jacksonville; Marjine Vandebergh, PhD, University of Antwerp; Arthur Toga, PhD, Keck School of Medicine of University of Southern California; Sandra Weintraub, PhD, Northwestern University; Dylan Wint, MD, Cleveland Clinic Las Vegas; Amy Wolf, BS, UCSF; Jennifer Yokoyama, PhD, UCSF.

Conflicts of Interest
BA has received research funding from the National Institutes of Health (NIH), the Centers for Disease Control and Prevention, the CIF Foundation, Alector, and Ionis; consulting fees from Acadia, Ionis, Sangamo, Gate, and Merck; and royalties from Wolters Kluwer. GSD reports no conflicts of interest that are directly relevant to this work. His research is supported by the NIH (K23AG064029, U01AG057195, U01NS120901, and U19AG032438), the Alzheimer’s Association, and the Chan Zuckerberg Initiative. He serves as a consultant for Parabon Nanolabs Inc; as a topic editor (dementia) for DynaMed (EBSCO); and as the clinical director of the Anti-NMDA Receptor Encephalitis Foundation Inc, Canada (uncompensated). He is the co–project principal investigator for a clinical trial in anti–N-methyl-D-aspartate (NMDA) receptor encephalitis, which receives support from Horizon Pharmaceuticals. He has developed educational materials for PeerView, and Continuing Education Inc. He owns stock in ANI Pharmaceuticals. His institution has received support from Eli Lilly for development and participation in an educational event promoting early diagnosis of symptomatic Alzheimer disease and in-kind contributions of radiotracer precursors for tau–positron emission tomography (PET) neuroimaging in studies of memory and aging (via Avid Radiopharmaceuticals, a wholly owned subsidiary of Eli Lilly). NG has participated, or is currently participating, in clinical trials of antidegenerative drugs sponsored by Bristol Myers Squibb, Eli Lilly and Avid Radiopharmaceuticals, Janssen Immunotherapy, Novartis, Pfizer, and Wyeth, as well as the Study of Nasal Insulin to Fight Forgetfulness (SNIFF) and the Anti-Amyloid Treatment in Asymptomatic Alzheimer’s Disease (A4) trial. She receives research support from the Tau Consortium and the Association for Frontotemporal Dementia and is funded by the NIH. IL’s research is supported by the NIH (R01AG038791-06A, U01NS100610, R25NS098999, U19AG063911-1, and 1R21NS114764-01A1), the Michael J Fox Foundation, the Parkinson’s Foundation, the Lewy Body Dementia Association, CurePSP, Roche, Abbvie, Biogen, Centogene, EIP Pharma, Biohaven Pharmaceuticals, Novartis, United BioPharma, and UCB. She is a member of the scientific advisory board for Amydis but does not receive funds and from the Rossy Progressive Supranuclear Palsy Program at the University of Toronto. She receives her salary from the University of California San Diego and as chief editor of Frontiers in Neurology. JCR is the site principal investigator for clinical trials sponsored by Eli Lilly and Eisai. He receives consulting fees from Roon. LVV is the site principal investigator for clinical trials sponsored by Biogen and has consulted for Retrotepo. All other authors declare no other conflicts of interest.

Multimedia Appendix 1
Visualization of raw battery percentage data binned by time of day (0=midnight; 23=11 PM) by disease severity group: (A) clinically normal, (B) prodromal frontotemporal lobar degeneration (FTLD), and (C) symptomatic FTLD.

References


Abbreviations

ALLFTD: Advancing Research and Treatment of Frontotemporal Lobar Degeneration (ARTFL study) and Longitudinal Evaluation of Familial Frontotemporal Dementia Subjects (LEFFTDS study) Longitudinal Frontotemporal Lobar Degeneration
API: application programming interface
ARTFL: Advancing Research and Treatment of Frontotemporal Lobar Degeneration
C9orfl72: chromosome 9 open reading frame 72
CDR+NACC FTLD: Clinical Dementia Rating Dementia Staging Instrument plus National Alzheimer’s Coordinating Center Frontotemporal Lobar Degeneration Behavior and Language Domains
FTLD: frontotemporal lobar degeneration
GRN: progranulin
LEFFTDS: Longitudinal Evaluation of Familial Frontotemporal Dementia Subjects
LME: linear mixed effects
MAPT: microtubule-associated protein tau
MRI: magnetic resonance imaging

©Emily W Paolillo, Kaitlin B Casaletto, Annie L Clark, Jack C Taylor, Hilary W Heuer, Amy B Wise, Sreya Dhanam, Mark Sanderson-Cimino, Rowan Saloner, Joel H Kramer, John Kornak, Walter Kremer, Leah Forsberg, Brian Appleby, Ece Bayram, Andrea Bozoki, Danielle Brushaber, R Ryan Darby, Gregory S Day, Bradford C Dickerson, Kimiko Domoto-Reilly, Fanny Elahi, Julie A Fields, Nupur Ghoshal, Neill Graff-Radford, Matthew G H Hall, Lawrence S Honig, Edward D Huey, Maria I Lapid, Irene Litvan, Ian R Mackenzie, Joseph C Masdeu, Mario F Mendez, Carly Mester, Toji Miyagawa, George Naasan, Belen Pascual, Peter Pressman, Eliana Marisa Ramos, Katherine P Rankin, Jessica Rexach, Julio C Rojas, Lawren VandeVrede, Bonnie Wong, Zbigniew K Wszolek, Bradley F Boeve, Howard J Rosen, Adam L Boxer, Adam M Staffaroni, ALLFTD Consortium. Originally published in JMIR Aging (https://aging.jmir.org), 26.06.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Home-Based Cognitive Intervention for Healthy Older Adults Through Asking Robots Questions: Randomized Controlled Trial

Seiki Tokunaga¹, PhD; Takuya Sekiguchi¹, PhD; Kumi Watanabe Miura¹, PhD; Hikaru Sugimoto¹, PhD; Masato S Abe¹², PhD; Kazuhiro Tamura¹, MEng; Taishiro Kishimoto¹³, MD, PhD; Takashi Kudo¹⁴, MD, PhD; Mihoko Otake-Matsuura¹, PhD

¹ Corresponding Author: Seiki Tokunaga, PhD

Abstract

Background: Asking questions is common in conversations, and while asking questions, we need to listen carefully to what others say and consider the perspective our questions adopt. However, difficulties persist in verifying the effect of asking questions on older adults’ cognitive function due to the lack of a standardized system for conducting experiments at participants’ homes.

Objective: This study examined the intervention effect of cognitive training moderated by robots on healthy older adults. A focus on the feasibility of the intervention at participants’ homes was also maintained. Feasibility was evaluated by considering both the dropout rate during the intervention and the number of questions posed to each participant during the experiment.

Methods: We conducted a randomized controlled trial with 81 adults older than 65 years. Participants were recruited through postal invitations and then randomized into 2 groups. The intervention group (n=40) received sessions where participants listened to photo-integrated stories and posed questions to the robots. The control group (n=41) received sessions where participants listened to photo-integrated stories and only thanked the robots for confirming participation. The participants participated in 12 dialogue sessions for 2-3 weeks. Scores of global cognitive functioning tests, recall tests, and verbal fluency tasks measured before and after the intervention were compared between the 2 groups.

Results: There was no significant intervention effect on the Telephone Interview for Cognitive Status-Japanese scores, recall tests, and verbal fluency tasks. Additionally, our study successfully concluded with no participant dropouts at follow-up, confirming the feasibility of our approach.

Conclusions: There was no statistically significant evidence indicating intervention benefits for cognitive functioning. Although the feasibility of home-based interventions was demonstrated, we identified areas for improvement in the future, such as setting up more efficient session themes. Further research is required to identify the effectiveness of an improved cognitive intervention involving the act of asking questions.

Trial Registration: University Hospital Medical Information Network Center UMIN000039489; https://center6.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000045027

JMIR Aging 2024;7:e47229 doi:10.2196/47229

Keywords

cognitive intervention; home-based experiment; robots; older adults; technology adoption; digital health

Introduction

The aging of the world’s population has led to a growing interest in maintaining a healthy lifestyle and enhancing the quality of life in later years. Previous studies have suggested that a healthy lifestyle can prevent or delay age-related cognitive decline [1,2].

Social interaction is a key component of a healthy lifestyle in later years. The social isolation attributed to social distancing during the COVID-19 pandemic may have particularly impacted older adults. For example, the impact of loneliness due to isolation on mental health is concerning [3].

Another concern is that reduced social participation can increase the risk of cognitive decline [4]. Conversations with a variety of people can be a trigger for understanding others’ perspectives and acquiring new information, which plays an important role in maintaining and improving older adults’ cognitive function [5]. Sharifian et al [6] showed that older adults with a higher proportion of family members in their social networks have less
contact with friends, which is negatively associated with their episodic memory. As the aforementioned study conjectured, contact with family members is usually restricted to obligatory tasks. In contrast, contact with friends is more likely to involve new conversations and information exchanges, which may be cognitively beneficial. Thus, cognitive maintenance and improvement mechanisms may be absent for older adults whose social participation is limited, such as those who only communicate with family members or caregivers or those who are completely isolated. The lives of older adults often involve factors that prevent social participation, such as the COVID-19 pandemic [7]. Therefore, methods to safely and remotely deliver cognitive training programs have been developed [8,9]. Dodge et al [8] have reported the effect of improving language-based executive function for older adults who have mild cognitive impairments through discussion intervention. However, previous studies still indicate a need for conversational human partners; hence, full automation of conversations needs further study.

Another issue in providing cognitive training for older adults is the limited available human resources for performance [10]. For example, a software agent that learns user characteristics, such as an intelligent assistant [11], could help older adults manage their health based on personal data collected automatically [12]. Furthermore, socially assistive robots may reduce the burden on caregivers to continuously monitor older adults who live alone with cognitive impairments and are at daily risk of various accidents [13].

Hence, our goal is to develop assistive robots that enable older adults’ remote participation in conversational cognitive training with the same degree of effectiveness as in-person social interaction for cognitive function. Home-based cognitively assistive robots aim to conduct cognitive training for older adults at home. A previous study has suggested that cognitively assistive robots have the potential to benefit older adults and society [14,15]; however, few studies have rigorously evaluated their benefits [16]. The challenge to overcome in promoting such research is the difficulty in controlling users’ characteristics related to speech. For example, depending on personality and familiarity with device use, the amount of conversation with the robot may differ from person to person. Therefore, the training effects provided by the robot cannot be accurately evaluated without controlling for such factors.

Similarly, we developed a conversational intervention program, Photo-Integrated Conversation Moderated by Robots (PICMOR), and examined its effect on healthy older adults’ cognitive function [17]. Briefly, the PICMOR program is a group conversation that uses photos taken by the participants beforehand. The program consists of 2 parts. First, the participants elaborate on the photos. Second, the participants receive questions about the photos from other participants and answer them. Each part has a time limit and is controlled by a robot facilitator. Notably, the questioning time for each participant is controlled by the robot. It has the function of encouraging participants who talk too much or too little to reduce or promote their speech as needed.

In a randomized controlled trial (RCT), we observed the beneficial effects of PICMOR on performance in a letter fluency task [17]. A follow-up experiment using multimodal magnetic resonance imaging provided candidate brain metrics that could be associated with the intervention effects on phonemic verbal fluency [18-20]. For instance, resting-state functional connectivity between the left inferior frontal gyrus, one of the most important brain regions for verbal fluency, and the right temporal pole, a semantic-related brain region, positively correlated with enhanced verbal fluency performance [18]. Moreover, we conducted another RCT using PICMOR and examined whether the intervention effects on verbal fluency varied as a function of neural states estimated from blood-based biomarkers, such as plasma neurofilament light chain [21]. The results showed that individuals with lower neurofilament light chain, indicating a relatively intact neuronal state, performed better in a category fluency task.

Despite these observed benefits, we could not accurately identify the components of this intervention program that contribute to the enhancement of verbal fluency [18]. This is because the intervention methodology included a variety of cognitive and mental activities, such as preparing a short presentation within a certain length of time, flexibly asking and answering questions among participants, intentionally storing and manipulating information to ask questions, and refraining from interrupting other participants’ utterances.

As merely developing methods to improve performance on specific tasks is not enough to improve cognitive function generally applicable to daily life, intervention strategies that bundle multiple components are being researched [22-25]. However, distinguishing these components to clarify the mechanisms underlying the intervention effects and develop more effective intervention methods is also important. In this study, we explored the effects of “asking questions” on cognitive functions among healthy older adults, assuming that it would be an important factor in verbal fluency enhancement.

This study conducted an RCT to collect evidence on the feasibility of asking questions to robots at home and its effect on the cognitive functions of healthy older adults. Our hypothesis is that cognitive function will improve in the intervention group compared to the control group. Additionally, the effects and future improvements of the intervention program are discussed.

Methods

Trial Design

This study used an RCT with a two-parallel-arm design and 1:1 allocation. All RCT procedures were conducted from February to November 2021 at the participants’ homes. Figure 1 presents the CONSORT (Consolidated Standards of Reporting Trials) flowchart of this trial. No participants dropped out of the intervention at follow-up.
Ethical Considerations
The studies involving human participants were reviewed and approved by the RIKEN Ethical Committee. All participants provided written informed consent to participate in the study. The study is registered with ClinicalTrials.gov (UMIN000039489).

Participants
Participants were community-dwelling older adults with subjective memory concerns living in an urban city (Wako-shi) in Japan. They were recruited through postal invitations. In total, 92 participants were screened for eligibility. The eligibility criteria for the trial were as follows: (1) age ≥65 years, (2) Telephone Interview for Cognitive Status-Japanese (TICS-J) score ≥33, and (3) complaints of cognitive concern. The exclusion criteria were as follows: (1) any neurological impairment known to affect the central nervous system, (2) any serious complicating disorder, (3) any history of serious head injury, (4) any disease or medication known to affect the central nervous system, (5) medical history of stroke, and (6) need for care. We defined the term “need for care” as certification of care needs levels or “support need levels” in the Japanese public long-term care insurance system, and participants were screened based on their self-report.

Intervention Design
In the intervention group, participants listened to stories and subsequently asked the dialogue robots as many questions as
possible during the experiment. In the control group, participants listened to the stories and subsequently offered short greetings to the robots as evidence of participation. Participants received 2-3 intervention sessions weekly; we arranged the schedules with participants individually to complete 12 dialogue sessions. The theme of each session was based on our previous studies [17,26]. The training and intervention procedure is described as follows: before the experiment, all the experimental devices were mailed to each participant. The system was preliminarily set up for each participant, eliminating the need for them to log in to the system. On the first day of the experiment, participants learned how to use the robot and tablets through Zoom. Then, a practice session was conducted via Zoom. In addition, 2 sessions were conducted with the administrator’s face hidden so that participants could become familiar with the devices. Finally, the participants were asked to answer questionnaires on the last day. For other experimental dates, participants received an intervention program.

The experimental procedure for both groups included listening to a story and subsequently asking questions to robots while looking at a photo. Both the photos and a summary of the older adults’ conversation (called a story) were preliminarily collected from 2 older adults [27]. For storytelling, the length of the story was adjusted to 30-40 seconds, referring to the logical memory task of the Wechsler Memory Scale-Revised (WMS-R) [28]. In the intervention group, participants posed as many questions as possible within 4 minutes. When the participants asked questions, the robots provided a plausible response from a list of approximately 550 responses collected beforehand. Contrastingly, in the control group, participants indicated their participation by simply saying “Thank you for the conversation” within 1 minute. Participants were required to push a switch before each utterance, as the pushing switch had a trigger function for the robots to activate question-answering mode. Further details are presented in Figure 2. Therefore, the difference between the intervention and control groups was whether questions were asked during the dialogue session.
For the experimental devices in our experimental setting, we used an original robot called Bono-06 [29] as a user interface for older adults (shown in Figure 3A). Bono-06 has 1 degree of freedom for nodding its head. Red, green, and blue full-color LEDs on its cheeks indicate the system’s status, such as whether it has successfully connected to a tablet. Additionally, a push switch on the chest allows intuitive interaction with older adults. In this study, the robots were designed to enable older adults to ask them questions by only pushing a switch. We also developed an original app that manages participants’ experimental schedules and displays photos and experimental time in dialogue sessions (shown in Figure 3B). The app was designed to display participants’ experimental schedules and run them automatically so that participants could participate in the experiment by turning on their tablets and robots at home without cumbersome operations, such as taps and swipes.
Finally, the system’s operation in home-based experiments was also considered [26]. A delivery and reporting function was implemented in our system so that the experiment administrator could remotely set up each experimental session as scheduled and observe the system report of each session, including the number of utterances and errors. In addition, the experiment report, including the transcribed audio, namely utterances for each dialogue session, was stored in a database on the cloud server. Thus, the experiment administrator could remotely monitor and download the results. In other words, the administrator monitored the experiment while the actual experiment was being conducted in the participants’ homes.

Outcome

Our primary concern was to examine the extent to which our intervention program improved cognitive function. The TICS-J [30], recall tests, and verbal fluency tasks [31] were assessed before and after the intervention as primary outcomes by well-trained psychologists.

The TICS-J, an 11-item cognitive test, was used to assess global cognition [30]. TICS-J included the immediate recall test, in which the participants were asked to recall a 10-word list. In addition to that, we also conducted a delayed recall test 5 minutes after the immediate recall test. The numbers of words recalled were the scores of these tests. For comparison with a
previous RCT [17], the results of the immediate recall test were also reported separately from the total TICS-J score.

Verbal fluency tests were conducted to assess verbal and executive control abilities. Two types of verbal fluency tasks were performed: letter and category [31]. In the letter fluency task, participants were required to produce as many words as possible, beginning with a given letter (“ka” in Japanese) within 1 minute. In the category fluency task, they were asked to produce as many words as possible belonging to a specific category (animals) within 1 minute. The number of words generated was the score for each task. All tests were conducted via telephone interviews.

As a secondary aim, we also investigated the intervention effect on the suboutcomes using questionnaires, including the World Health Organization’s 26-item Quality of Life questionnaire [32], the Japanese version of the Geriatric Depression Scale (short form) [33], and the Tokyo Metropolitan Institute of Gerontology-Index of Competence [34], to assess quality of life, depression symptoms, and functional capacity.

We also examined the factors of our intervention program that should be improved in the future to increase its effectiveness, which will be mainly reported in the Ancillary Analysis section. First, we counted the total number of utterances to measure the extent to which our intervention prompted participants to speak. In this study, the number of turns taken by a participant in a conversation with the robot was defined as the number of utterances. Second, we related it to participants’ cognitive test scores to examine how much the intervention worked differently, depending on their cognitive functions. Third, we investigated the extent to which participants’ use of digital devices affected their total amount of utterances. This was achieved by asking participants to answer a 4-item questionnaire about the frequency of their use of PCs, emails, smartphones, and flip phones in their daily lives (1: usually; 2: sometimes; 3: rare; 4: never).

The reason for the distinction between smartphones and flip phones lies in the history of the cell phone market in Japan [35]. Although flip phones used to be popular in Japan, smartphone use began to exceed that of flip phones in 2013. The difference in the time of popularization between the two could result in different user demographics; in other words, those who still frequently use flip phones may be less likely to switch to a new device compared to those who use smartphones. This has particular implications among older adults, the target population of our study. In fact, in 2020, the percentage of Japanese people in their sixties using flip phones was about 26%, while that of Japanese people in their twenties was about 12%. Considering the possibility that these differences between both users might affect their attitude toward the device used in our intervention, we decided to ask them separately about their use.

Randomization Implementation

Stratified block randomization with a 1:1 allocation was implemented. Participants were stratified into male and female groups and then sorted based on total TICS-J scores. Subsequently, blocks of size 2 were created and randomized. The coding was performed in R 4.3.2 (R Foundation for Statistical Computing). This determined which participants belonged to the intervention or control group. The experimenter assigned participants to the 2 groups based on this result. The person who conducted the randomization was different from the experimenter and had no information other than IDs, TICS-J scores, and gender at the time of randomization. The assessors were blinded to the allocation results.

Statistical Analysis

As explained in the Outcome and Estimation section, we used linear mixed models with random intercepts to examine the effects of the intervention on cognitive function. The models included total TICS-J scores, immediate and delayed recall test scores, letter fluency test scores, and category fluency scores as outcome variables, with time (1: end point; 0: baseline), group (1: intervention group; 0: control group), and their interaction terms as independent variables. We interpreted the regression coefficients associated with the interaction terms as the degree of the intervention effects. We also reported the sizes of intervention effects measured by $r^2$ [36].

For the ancillary analysis, we applied linear mixed models with random intercepts to the intervention group, which included cognitive function scores as outcome variables; time, the number of utterances, and their interaction terms as explanatory variables; and gender, age, and education as control variables. For these models, we reported regression coefficients associated with the number of utterances to understand the relationship between cognitive function scores at baseline and the number of utterances; we also reported regression coefficients associated with the interaction term to understand the relationship between the number of utterances and change in scores before and after the intervention. In addition, we reported the relationship between participants’ digital device use in the intervention group and their total number of utterances using 2-tailed $t$ tests.

All analyses were performed using R. To implement the linear mixed models, the lmer function in the R package (lme4) was used [37].

Results

Baseline Data

A total of 40 participants in the intervention group and 41 in the control group underwent cognitive testing, both at baseline and end point, and were included in the analysis. Table 1 shows the baseline characteristics of the participants. For all demographic and cognitive variables (Table 1 and Table 2), there were no major differences between the intervention and control groups, namely, participant attributes were balanced at baseline.
### Table 1. Baseline characteristics of the intervention and control groups (N=81).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Intervention (n=40)</th>
<th>Control (n=41)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (year), mean (SD)</td>
<td>73.9 (3.8)</td>
<td>74.0 (4.1)</td>
</tr>
<tr>
<td>Gender (female), n (%)</td>
<td>24 (60)</td>
<td>25 (61)</td>
</tr>
<tr>
<td>Education (≥13 years), n (%)</td>
<td>19 (48)</td>
<td>30 (73)</td>
</tr>
<tr>
<td>WHO QOL26 questionnaire, mean (SD)</td>
<td>3.68 (0.38)</td>
<td>3.67 (0.41)</td>
</tr>
<tr>
<td>GDS-15c, mean (SD)</td>
<td>2.17 (2.00)</td>
<td>2.05 (2.32)</td>
</tr>
<tr>
<td><strong>TMIG-IC</strong> d, mean (SD)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total score</td>
<td>11.93 (1.1)</td>
</tr>
<tr>
<td></td>
<td>Instrumental activity of daily living</td>
<td>5.00 (0.0)</td>
</tr>
<tr>
<td></td>
<td>Intellectual activity</td>
<td>3.62 (0.7)</td>
</tr>
<tr>
<td></td>
<td>Social role</td>
<td>3.30 (0.9)</td>
</tr>
</tbody>
</table>

*a*WHO: World Health Organization.

*b*QOL26: Quality of Life 26-item. For the WHO QOL26, 2 participants who selected multiple items were excluded from the intervention group.

*c*GDS-15: Geriatric Depression Scale-15.

*d*TMIG-IC: Tokyo Metropolitan Institute of Gerontology Index of Competence.

### Table 2. Comparison of cognitive test scores at baseline and end point between the intervention and control groups.

<table>
<thead>
<tr>
<th>Cognitive test</th>
<th>Intervention (n=40)</th>
<th>Control (n=41)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline, mean (SD)</td>
<td>End point, mean (SD)</td>
</tr>
<tr>
<td>TICS-J</td>
<td>36.30 (2.03)</td>
<td>36.80 (2.57)</td>
</tr>
<tr>
<td>Category fluency</td>
<td>15.47 (4.49)</td>
<td>17.27 (5.09)</td>
</tr>
<tr>
<td>Letter fluency</td>
<td>13.93 (3.55)</td>
<td>14.43 (3.97)</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>7.20 (1.40)</td>
<td>7.90 (1.39)</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>6.30 (1.94)</td>
<td>7.17 (1.82)</td>
</tr>
</tbody>
</table>


### Outcome and Estimation

Table 3 shows the results of the linear mixed models used to examine the intervention effects on cognitive function. No significant intervention effects were found for the total TICS-J score, immediate and delayed recall, and verbal fluency. The regression coefficient associated with the time-group interaction term for the category fluency test score was the largest among the study outcomes (1.361), but the effect was not significant ($P=.09; \, \text{effect sizes of } \hat{f}^2<0.01 \text{ for all outcomes}$). We found no significant intervention effects on any of the secondary outcomes. Table 4 shows the conversation theme of each session and the descriptive statistics for the number of participants’ utterances therein.
<table>
<thead>
<tr>
<th>Category</th>
<th>Time</th>
<th>SE</th>
<th>t (df) (^a)</th>
<th>P value</th>
<th>Group</th>
<th>Coefficients (95%CI)</th>
<th>SE</th>
<th>t (df) (^a)</th>
<th>P value</th>
<th>Time \times group</th>
<th>Coefficients (95%CI)</th>
<th>SE</th>
<th>t (df) (^a)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category fluency</td>
<td>0.44</td>
<td>0.56</td>
<td>0.79 (79)</td>
<td>.43</td>
<td>0.34</td>
<td>(-0.2 to 1.33)</td>
<td>0.68</td>
<td>-0.4 (109.1)</td>
<td>.69</td>
<td>0.4 (1.16 to 1.96)</td>
<td>0.8 (196)</td>
<td>0.8</td>
<td>-0.3 (79)</td>
<td>.76</td>
</tr>
<tr>
<td>Letter fluency</td>
<td>0.10</td>
<td>0.56</td>
<td>0.17 (79)</td>
<td>.86</td>
<td>0.86</td>
<td>(-0.2 to 1.33)</td>
<td>0.68</td>
<td>-0.4 (109.1)</td>
<td>.69</td>
<td>0.4 (1.16 to 1.96)</td>
<td>0.8 (196)</td>
<td>0.8</td>
<td>-0.3 (79)</td>
<td>.76</td>
</tr>
<tr>
<td>Immediate recall</td>
<td>0.80</td>
<td>0.24</td>
<td>3.39 (79)</td>
<td>.001</td>
<td>0.84</td>
<td>(-0.22 to 1.01)</td>
<td>1.12</td>
<td>0.4 (133.8)</td>
<td>.21</td>
<td>-0.1 (-0.77 to 0.56)</td>
<td>0.3 (79)</td>
<td>0.3</td>
<td>-0.3 (79)</td>
<td>.76</td>
</tr>
<tr>
<td>Delayed recall</td>
<td>0.98</td>
<td>0.29</td>
<td>3.38 (79)</td>
<td>.001</td>
<td>0.38</td>
<td>(-0.42 to 1.07)</td>
<td>0.85</td>
<td>0.4 (133.5)</td>
<td>.40</td>
<td>-0.1 (-0.9 to 0.70)</td>
<td>0.4 (79)</td>
<td>0.4</td>
<td>-0.25 (79)</td>
<td>.81</td>
</tr>
</tbody>
</table>

\(^a\)Satterthwaite degree of freedom.

\(^b\)TICS-J: Telephone Interview for Cognitive Status-Japanese.

Table. Descriptive statistics of the number of utterances per session. The participant count in the intervention group is 40. Mean, SD, minimum, and maximum are at the participant level.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Total, n</th>
<th>Mean (SD)</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Favorite things</td>
<td>291</td>
<td>7.28 (2.20)</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>2. Neighborhood landmarks</td>
<td>344</td>
<td>8.60 (1.85)</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>3. I try to get off the train at a station that I seldom use</td>
<td>346</td>
<td>8.65 (2.05)</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>4. Favorite foods</td>
<td>346</td>
<td>8.65 (1.70)</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>5. For my health</td>
<td>335</td>
<td>8.38 (1.23)</td>
<td>5</td>
<td>10</td>
</tr>
<tr>
<td>6. Found on a 10-minute walk</td>
<td>363</td>
<td>9.07 (1.94)</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>7. Saving energy</td>
<td>317</td>
<td>7.92 (1.67)</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>8. Funny stories and mistakes</td>
<td>328</td>
<td>8.20 (1.92)</td>
<td>4</td>
<td>11</td>
</tr>
<tr>
<td>9. Things to get rid of</td>
<td>323</td>
<td>8.07 (1.83)</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>10. Tips for daily living</td>
<td>353</td>
<td>8.82 (1.80)</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>11. Feeling the season</td>
<td>359</td>
<td>8.97 (1.75)</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td>12. Starting something new</td>
<td>366</td>
<td>9.15 (1.72)</td>
<td>5</td>
<td>12</td>
</tr>
</tbody>
</table>

Ancillary Analysis

The average total number of utterances of the participants in the intervention group was 101.78 (SD 14.72). The number of utterances was positively correlated with a higher letter fluency score at baseline (letter fluency: B=0.11, SE 0.04; P=.01), while no significant associations were found for the other outcomes (TICS-J: B=0.03, SE 0.03; P=.30; category fluency: B=0.10, SE 0.06; P=.07; immediate recall: B=0.01, SE 0.02; P=.64; delayed recall: B=0.01, SE 0.02; P=.55).

There was no significant association between the number of utterances and change in scores for any of the outcomes (TICS-J: B=-0.03, SE 0.03; P=.31; category fluency: B=-0.01, SE 0.04; P=.79; letter fluency: B=0.02, SE 0.03; P=.63; immediate recall: B=-0.003, SE 0.02; P=.85; delayed recall: B=-0.01, SE 0.02; P=.76).

Figure 4 shows the relationship between the digital device use of participants in the intervention group and their total number of utterances. In Figure 4A, we notice that participants who never used a computer had fewer utterances. In fact, the average number of utterances for participants who chose “4: never” was...
86.50, compared to 108.32 for participants who chose options other than “4: never.” This difference was significant ($P<.001$). In Figure 4B, we observe that participants who usually used email had more utterances. In fact, the average number of utterances for participants who chose “1: usually” was 106.58, compared to 92.8 for participants who chose options other than “1: usually.” This difference was significant ($P=.005$). No significant associations were found between the frequency of smartphone use and flip phone use and the number of utterances (Figure 4C and 4D). This result persisted in regression analyses, even after controlling for age, gender, and education years. Associate regression coefficients were 19.69 (SE 4.15; $P<.001$) for computer use, 12.06 (SE 4.44; $P=.01$) for email use, 3.74 (SE 5.39; $P=.49$) for smartphone use, and $-2.89$ (SE 6.75; $P=.67$) for flip phone use.

**Figure 4.** Box plots for the relationship between the frequency of device use by participants in the intervention group (horizontal axis) and the total amount of their utterances (vertical axis). Boxes represent IQRs. White circles indicate the mean values. Black lines in boxes indicate the median values. Observations outside the first (third) quartile (i.e., outside of $1.5 \times$ IQR) are indicated by black circles. No participant selected “4: never” for C and D.
Discussion

Overview of This Study

This paper presented the intervention effect of asking questions to improve cognitive function in healthy older adults. The intervention involved an RCT conducted at participants’ homes. There were no significant intervention effects on scores of TICS-J, recall tests, and verbal fluency tasks. Notably, the feasibility of the intervention was confirmed, as all participants were able to ask at least 1 question in every session, with no participants dropping out.

Principal Results

This study identified no significant intervention effect in category fluency task scores and letter fluency task scores. In both types of verbal fluency tasks, the number of words produced per unit of time is commonly used as a behavioral index. However, the mechanisms of word production are supposedly quite different. The category fluency test requires retrieval of content included in a given semantic category, which involves access to semantic memory. This mechanism helps one use existing links between related concepts, such as those between the categorical label and its contents and among associated category members [38]. In contrast, the letter fluency task demands retrieval from a phonemic category in which the association of semantically related words should be suppressed [38]. This strategy depends on an effortful exploration of lexical systems. To ask questions, one must understand the meanings of others’ utterances and identify what is unclear with reference to one’s knowledge. A series of these processes would inevitably involve access to semantic memory. Nonetheless, no intervention effect was found on category verbal fluency and letter fluency. This result may have been influenced by the short duration of the intervention. Other studies have involved longer experimental periods, lasting 6 [8] and 12 weeks [17]. Therefore, further studies are needed to clarify the effectiveness of asking questions, for example, by setting up a program with a longer intervention period. Another reason why the results did not show any significant intervention effects may pertain to various elements included in the conversations. We solely focused on asking questions and did not incorporate other elements into the intervention. Older adults who are more inclined to compare similarities and differences between one’s thoughts and those of others, may be effective in strengthening future interventions. This study has some strengths. One strength is that we presented a technical framework for examining the impact of “asking questions,” an important factor for conversation-based interventions for cognitive function, which was lacking in our previous studies [17,21].

Another strength of this study is that we successfully conducted home-based interventions while most intervention studies, including our previous ones [17,21], have been conducted on-site. Notably, this study demonstrated that home-based interventions are feasible even though some participants were accustomed to the system through repeated session participation. A more extended intervention period may compensate for the disadvantage owing to the lower frequencies of device use.

However, in terms of the social implementation of intervention programs, such a view may be too optimistic. This is because users may want to discontinue use of the program before becoming accustomed to it. From this perspective, we need to continue improving usability and developing appropriate evaluation methods, keeping in mind that our target population is older adults [39]. For example, it would be helpful to investigate which individuals, among those who use less frequently digital devices at baseline, are more likely to experience higher intervention effectiveness.

Furthermore, it would be important to have a perspective on what kind of interface to provide and how to personalize it according to the user’s preferences or personality as well as the characteristics of the user’s daily conversations [40], to keep them motivated to continue with the intervention program. This could lead to an overall increase in the effectiveness of the intervention, benefiting even those who are already accustomed to using the device.

Strengths

This study has some strengths. One strength is that we presented a technical framework for examining the impact of “asking questions,” an important factor for conversation-based interventions for cognitive function, which was lacking in our previous studies [17,21].

Another strength of this study is that we successfully conducted home-based interventions while most intervention studies, including our previous ones [17,21], have been conducted on-site. Notably, this study demonstrated that home-based interventions are feasible even though some participants were unfamiliar with digital devices.

Finally, a strength of our study is that the system is fully automated after human-assisted training, requiring fewer human resources compared to previous methods. The impact of automation on participants’ satisfaction needs further investigation.

Limitations

This study has several limitations. First, this study is limited by the diversity and sample size of the participants. Currently, this program only supports Japanese people; therefore, this study was conducted in an urban city in Japan for feasibility. Second, this study was conducted during the COVID-19 pandemic; hence, the research was designed to reduce communication as much as possible to keep the participants safe. To achieve this,
we did not collect magnetic resonance imaging data from participants, unlike our previous intervention study [17]. Therefore, from a neural perspective, this study could not obtain useful information about the impact of asking questions on improving the intervention effect on cognitive function. The importance of asking questions in conversation-based intervention programs should be reevaluated in future research after accounting for these issues.

Finally, we did not sufficiently examine the effects of session topics and their order on the number of participants’ utterances. The topics suitable for improving cognitive function through conversation must be neither so challenging that participants cannot generate questions nor so easy that they do not train cognitive function. There has already been a study that identified the characteristics of utterances by older adults with higher cognitive functioning in group conversations, and then, identified conversation topics in which such utterances are likely to be observed [41]. However, this has not been examined in terms of encouraging participants to ask questions and enriching the conversation. Future studies should consider setting up efficient session themes for improving cognitive function.

Conclusions

This study evaluated the possible improvements associated with introducing a dialogue-based robot in cognitive interventions, aiming to verify the training effects of asking questions in healthy older adults.

We did not observe any significant differences in global cognition between the 2 groups. The feasibility of our study was identified by (1) no loss in the intervention and (2) all participants asking at least 1 question in every session. We also recommend improvements to the intervention program, such as setting up more efficient session themes for cognitive training. This study has provided future directions for cognitive training studies of older adults at home.

Acknowledgments

We are grateful to all the study participants and staff. We would especially like to thank Sachiko Iwata and Kumiko Umemoto for their support. This research was partially supported by Japan Society for the Promotion of Science (JSPS) KAKENHI (grants JP18KT0035, JP19H01138, JP20H05022, JP20H05574, JP20K19471, JP22H04872, and JP22H00544) and the Japan Science and Technology Agency (grants JPMJCR20G1, JPMJPF2101, and JPMJMS2237).

Conflicts of Interest

None declared.

References


Abbreviations
- CONSORT: Consolidated Standards of Reporting Trials
- PICMOR: Photo-Integrated Conversation Moderated by Robots
- RCT: randomized controlled trial
- TICS-J: Telephone Interview for Cognitive Status-Japanese
- WMS-R: Wechsler Memory Scale-Revised

Edited by J Wang, Y Jiang; submitted 19.04.23; peer-reviewed by D Magistro, R Boumans; revised version received 30.01.24; accepted 30.01.24; published 22.04.24.

Please cite as:
Tokunaga S, Sekiguchi T, Watanabe Miura K, Sugimoto H, S Abe M, Tamura K, Kishimoto T, Kudo T, Otake-Matsuura M. Home-Based Cognitive Intervention for Healthy Older Adults Through Asking Robots Questions: Randomized Controlled Trial. JMIR Aging 2024;7:e47229
URL: https://aging.jmir.org/2024/1/e47229
doi:10.2196/47229

© Seiki Tokunaga, Takuya Sekiguchi, Kumi Watanabe Miura, Hikaru Sugimoto, Masato S Abe, Kazuhiro Tamura, Taishiro Kishimoto, Takashi Kudo, Mihoko Otake-Matsuura. Originally published in JMIR Aging (https://aging.jmir.org), 22.4.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original work on https://aging.jmir.org, as well as this copyright and license information must be included.
A Serious Game for Cognitive Stimulation of Older People With Mild Cognitive Impairment: Design and Pilot Usability Study

Juan Francisco Ortega Morán, ENG; J Blas Pagador, PhD; Vicente Gilete Preciado, PsyD; José Luis Moyano-Cuevas, ENG; Trinidad Rodríguez Domínguez, PhD; Marta Santurino Muñoz, BSN; Francisco M Sánchez Margallo, PhD

Corresponding Author: Juan Francisco Ortega Morán, ENG

Abstract

Background: Cognitive stimulation of older people helps prevent, and even treat, age-related diseases, such as mild cognitive impairment. Playing games reduces the probability of experiencing this pathology, which is related to the loss of the ability to carry out some instrumental activities of daily living.

Objective: This work describes the design and development of a serious game for the cognitive stimulation of older people, with exercises related to the daily life task of shopping. A pilot study for its preliminary usability validation is also presented.

Methods: The designed serious game includes 4 exercises consisting of shopping in a hypermarket, ordering products, making payments, and organizing the purchase, thus dealing with the most frequent cognitive problems of older people associated with episodic declarative memory, naming, calculation, and organization, respectively.

Results: A total of 19 older people participated in the pilot study for the usability validation of the serious game. They indicated that they like the aesthetic and interesting topic of the game. They reported that it provides a high level of entertainment and could be useful in daily life for mental stimulation. The participants found the serious game to be intuitive, but the ease of use and readability of the instructions could be improved.

Conclusions: This study suggests that the innovative serious game developed could be accepted by older people for their cognitive stimulation to prevent or treat mild cognitive impairment, although a long-term intervention study should be performed as future work. Its ecological validity design, with everyday tasks, adaptable levels of difficulty, and motivational mechanisms, is a differentiating factor compared to similar serious games.

JMIR Aging 2024;7:e41437 doi:10.2196/41437

KEYWORDS
serious game; mild cognitive impairment; cognitive stimulation; design; pilot study; older people; activities of daily living; shopping; ecological validity

Introduction

The increase in life expectancy means that we can live with age-related pathologies, such as dementia. However, the prevalence of this pathology is clearly on the rise, and there are currently about 55 million people affected worldwide [1] and around 11 million for the broad European region [2]. This acquired syndrome is characterized by a progressive deterioration of cognitive functions, which influences activities of daily living (ADL) and decreases the level of independence of individuals. For this reason, dementia is classified into grades, where simple activities can be performed in the mild stage but the severity of symptoms increases as the impairment progresses [2-4]. Mild cognitive impairment (MCI) is the term used for individuals whose cognitive changes fall between those of aging and early dementia and is considered a precursor to early dementia in around 30% of cases; thus, those with MCI are much more likely to progress to this type or level of dementia, especially in older people [4]. Treatment usually includes drugs that aim to reduce or delay cognitive, psychological, and behavioral symptoms [5], but nonpharmacological treatments are also very important. These include reality orientation, reminiscence and validation therapies, or cognitive stimulation [6].

Cognitive stimulation includes techniques that focus on treating cognitive aspects through activities that allow them to be worked on globally and simultaneously [6,7]. Cognitive and reminiscence activities, multisensory activities, as well as those that work on social aspects are used, along with group activities that facilitate integration and social participation among users [8]. In fact, several studies confirm an improvement in different aspects of the cognitive sphere and in the quality of life of people with dementia who have participated in this type of
therapies [9,10]. We also found improvements in mood and ADL maintenance [11].

Currently, information and communication technology (ICT) instruments and tools are incorporated into nonpharmacological treatments. These technological devices facilitate performance, which can lead to a greater sense of self-efficacy and improve one’s own perception of functionality, as well as reduce the burden of caregivers. Some of the devices that are commonly used with people with dementia are tele-assistance, devices that improve cognitive functions, and robotics [12]. Among them, the most widely used are those that can be connected to the internet (tablets, mobile phones, computers, and video game consoles), which provides a wide variety of resources that allow the diversification of therapies and increase motivation and adherence to treatment [8,13]. The use of ICTs as therapeutic tools requires a previous study of the person’s abilities and skills for their use and management, to avoid feelings of frustration that can lead to discouragement from therapy or rejection of the device [9]. Therefore, it is essential to gradually approach the device and its applications, looking for intuitive tasks that the user can easily carry out independently [9,13].

Serious games are ICT games whose main objective is to give a therapeutic or diagnosis value to the playful action of the games. They have been tested in different areas of intervention in pathologies with cognitive impairment [14,15], particularly dementia [16]. Through serious games, we can work to delay deficits, increase autonomy and relationships with their social environment, and improve the quality of life of people with dementia [17]. For this intervention, the occupational therapist may include in their individual intervention plans games to work on physical, cognitive, and social aspects. The literature reveals that there are already a number of well-established serious games that could improve older people’s cognitive health. Some examples of these are focused on sport simulations, quizzes with text and images, music tools, arithmetic and reading calculations, etc [17]. Others instead use a fantasy world setting, minigames, puzzles, or scenarios to cope with stressful or negative situations [18]. However, few serious games have been developed that incorporate tasks focused on ADL, such as cooking [19]. Once cognitive impairment becomes perceptible in ADL, the disease has usually progressed, so the analysis of ADL performance would help to assess cognitive status during the course of cognitive impairment [19]. In this sense, serious games can be a valuable tool for achieving this in a motivating and enjoyable way for older people with ecological scenarios.

In this study, we describe the design and development process of a serious game developed for the cognitive stimulation of older people, with exercises related to the daily life task of shopping. We also present the results of a usability validation from a pilot study of the game.

Methods

Game Development Process and Design Requirements

In the process of designing and developing this serious game, researchers with a bioengineering profile, game designers and developers (computer and telecommunication engineers), and health care professionals (neuropsychologists, psychologists, and occupational therapists) participated to fulfill the essential collaboration of an interdisciplinary team [20].

First, 2 focus groups were performed with 11 psychologists and 6 occupational therapists from Extremadura, Spain, to identify the needs, limitations, and motivations of older people to use cognitive stimulation programs. The selection criteria of participants in both focus groups were the geographical location (working in rural and urban areas) and the kind of institutions in which they work (public and private). The focus groups were conducted by 2 engineers with extensive experience in conducting interviews. After the presentation of the study and the objectives, a previously created guide with topics of interest and questions was followed to conduct the session in a semistructured way. With the participants’ consent, all conversations were recorded for subsequent transcription and analysis. NVivo 2017 (QSR International) software was used to carry out this analysis to facilitate the drawing of conclusions. Four thematic areas were identified: (1) the most frequent cognitive problems of older people are mainly focused on memory loss, disorientation, difficulty in performing executive tasks, or difficulty in concentrating; (2) motivation is the fundamental element for the success of new training exercises; (3) technological barriers are mainly related to interface design problems and cultural level; and (4) the low degree of awareness of older people regarding the importance of leading an active life. This served as a basis for designing the new tool to promote the cognitive training that directly stimulates memory and executive tasks, as well as orientation and concentration skills.

In the implementation of this serious game for Android OS, Unity 3D (C#, Visual Studio 2017) was used, following the design recommendations established in previous studies [21]. In this sense, the game interface was designed while taking into account that the target audience is older people and adapted to their needs, including the minimum necessary information in a clear and concise way, which allows the older people to understand the objective of the game. Regarding user interaction, the game was designed to be used on a tablet instead of the standard keyboard and mouse to provide a more natural interaction and facilitate the acceptance by older people [22].

Two fundamental characteristics were taken into account in the design of this serious game. First, a shopping task with high ecological validity was included, that is, a day-to-day activity of older people. In this way, a greater interest and acceptance by older people should be achieved, due to the high utility of this tool in their daily lives. The concept of ecological validity is determined by the degree of representativeness, that the game is represented in a form and context that correspond to its occurrence in everyday life, as well as by the level of generalization that the results are able to explain similar tasks in everyday life [23]. Second, the difficulty level of the game was customized by professionals according to the needs of each older person, as detailed in the description of each game in the following sections.

An additional characteristic included in the design of the game was that immediate feedback is provided to the user at the end of each game, showing the attempts, hits, and failures made.
after the user performance. Moreover, positive feedback with encouragement messages is showed to the user during the completion of the game, such as “Come on, you’re about to get it,” “You almost got it. Try again,” “Surely you have the name on the tip of your tongue,” and “Don’t worry. Let’s go with another product.”

**Game Description**

**First Exercise: Shopping in a Hypermarket**

Episodic declarative memory is the neurocognitive function that is affected the earliest in Alzheimer disease due to the initial involvement of the hippocampal formation in the medial temporal lobe, so the stimulation of episodic declarative memory in people with MCI who are affected by mnesic impairment is of utmost importance.

This exercise is designed with the specific aim of stimulating verbal intentional episodic declarative memory at the level of the 3 mnesic processes of encoding or fixation, consolidation or storage, and recall or retrieval of information, thus consisting of 3 phases:

- **1.1 Learning Phase:** memorization of a shopping list
- **1.2 Interference Phase:** preparation of the money to pay for the purchase
- **1.3 Recall Phase:** purchase of products from the shopping list

The level of difficulty can be adjusted by the number of products on the shopping list, the number of product categories, the memorization and interference times, the number of attempts, or the total allowed time.

**Second Exercise: Ordering Products in a Hypermarket**

On the one hand, expressive language impairment manifests itself very frequently and early in the form of anomia, a naming deficit that consists of the difficulty to recall the names of objects. Due to the great frustration caused by this early expressive language difficulty for people with MCI, it is essential to include naming stimulation tasks within the global stimulation of expressive language in cognitive stimulation programs.

On the other hand, complex visual gnosias, in which the visual recognition and identification of objects is hindered by a modification of the characteristics of the objects’ images, are one of the cognitive functions that are affected in the early stages of several primary cortical degenerative dementias, hence the importance of their stimulation.

This exercise is designed with the specific aim of stimulating language at the level of naming and complex visual gnosias in 2 phases:

- **2.1 Phase 1:** hypermarket product naming
- **2.2 Phase 2:** visual recognition and identification of hypermarket products

The difficulty can be customized by adjusting the product familiarity percentage, the number of products that appears, whether or not to provide help to recognize the products, or the allowed time.

**Third Exercise: Making Payments**

Executive attentional control processes are affected early in several types of primary and secondary dementias such as subcortical vascular dementia. Working memory, a key element of executive attentional control, is the ability to maintain and manipulate information in on-going cognitive activity, such as money management for shopping. It is therefore very important to place great emphasis on the stimulation of these cognitive processes to maintain the highest possible level of functional independence of each person.

This exercise consists of 3 phases and is designed with the specific aim of stimulating the working memory through calculation tasks by simulating purchase payments:

- **3.1 Phase 1:** payment of the purchase
- **3.2 Phase 2:** check the change (money returned)
- **3.3 Phase 3:** check the price of products on the purchase receipt with the catalog of offers

Professionals can adapt the difficulty by adjusting whether or not to include decimals in the amount to be paid or the change, the number of attempts, the allowed time, the number of correct banknotes or coins in the change, the number of products, pages and sections in the catalog of offers, or the rate of products with erroneous price.

**Fourth Exercise: Organizing the Purchase**

Executive functions, such as planning and organization, reasoning, cognitive flexibility, or monitoring when problem-solving, are crucial functions for a good performance of any adult in advanced and instrumental ADL. In the context of cognitive treatments for MCI, the stimulation of executive functions is a compulsory subject given their close interdependence with the maintenance of a high level of functional independence and personal autonomy in ADL.

This exercise is designed with the specific objective of stimulating this executive function of organization, as well as abstract reasoning, performance monitoring, visual gnosias, semantic memory, and visuospatial function.

The only phase of this exercise consists of the arrangement of the purchase products in the corresponding rooms of the house.

In this case, the adaptation of the difficulty can be achieved by adjusting the number of products on the list and the maximum time allowed to complete the task.

**Acceptance and Usability Study**

A preliminary validation of the serious game was carried out by 4 participants from the Association of Friends of the Minimally Invasive Surgery Centre in Cáceres, Spain; 5 participants from senior centers in Castelo Branco, Portugal; and 10 participants who attended the FEHISPOR fair held in Badajoz, Spain. The nonprobability intentional sampling technique was used to conduct the recruitment of participants, with inclusion criteria of older people aged 60-80 years with MCI. Through this method, older people from those attending various demonstrations of the game at the aforementioned events and venues were recruited to voluntarily participate in the study.
Participants tested the game until all exercises were completed and then filled in a questionnaire. The questionnaire consisted of questions scored on a 5-value Likert scale (1=lowest value, 5=highest value) about the acceptance and usability of the game, assessing (1) the user’s perception and degree of satisfaction when using the game and (2) the design and layout of the game. A descriptive analysis with the average values of the answers provided by the participants was performed.

**Ethical Considerations**

Both ethical approval and written participants consent were waived for this study because they were not within the scope of Law 14/2007 of 3rd July on Biomedical Research, due to this study not involving any invasive procedure on the participants. Only verbal informed consent was considered sufficient. All data have been deidentified, and there was no compensation for participation.

**Results**

**Game Implementation**

**First Exercise: Shopping in a Hypermarket**

This exercise consists of 3 phases.

In **1.1 Learning Phase**, the user must learn a series of products from a shopping list in 3 subphases (Figure 1): (1) organization of the list by categorization, since the structure of the information facilitates its deep encoding; (2) identification of specific characteristics that differentiate the products, since better fixation of information with clues optimizes recall; and (3) intentional memorization of the list in a specific time frame.
Figure 1. The first exercise stimulates episodic declarative memory with 3 mnesic processes: encoding, consolidation, and recall. The user must learn a series of products from a shopping list.

In 1.2 Interference Phase, the task simulates the preparation of the wallet with the money to pay for the purchase. It activates the processes of attentional control at the level of selective attention to choose the correct banknotes or coins, working memory to sum the money, and monitoring of the execution to avoid errors.

In 1.3 Recall Phase, the user is presented with a map of a hypermarket where the sections are marked out, and the user has to fill a shopping basket with those products from the shopping list studied in the 1.1 Learning Phase among distracting products. The aim is to stimulate the mnesic process of recalling short-term episodic declarative memory.

Second Exercise: Ordering Products in a Hypermarket
This exercise consists of 2 phases.
In 2.1 Phase 1, the user has to write the names of the hypermarket products that appear in the images (Figure 2).
Figure 2. The second exercise (2.1) stimulates expressive language. The user has to write the names of the products that appear in the images.

In 2.2 Phase 2, the user has to recognize and identify the hypermarket products through their distorted images and subsequently name these products (Figure 3). The emphasis is placed on the stimulation of complex visual gnosias, making it difficult to recognize and identify the objects to be named.
Figure 3. The second exercise (2.2) stimulates complex visual gnosis. The user has to recognize and write the names of the products through their distorted images.

Third Exercise: Making Payments

This exercise consists of 3 phases.

In 3.1 Phase 1, the user has to use banknotes or coins to make the exact payment of the purchase (Figure 4).
Figure 4. The third exercise stimulates executive attentional control. The user has to use banknotes or coins to make the exact payment of the purchase.

In 3.2 Phase 2, the user must check whether the change (money returned) received is correct, and if not, he or she must select the banknotes or coins necessary to make the change amount correct.

In 3.3 Phase 3, the user has to check the price charged for each product on the purchase receipt to verify whether it corresponds to the price for that product in the catalog, and in the case of error, to mark on the receipt the products for which the amount on the receipt is erroneous.

Fourth Exercise: Organizing the Purchase

In this exercise, the user must arrange the products of a purchase into the rooms of a house, whose floor plan appears on the screen and consists of the following spaces: kitchen, terrace, bathroom, pantry, living room, and bedroom (Figure 5). Within each of these house spaces, the correct storage location is flexible according to each product.
Figure 5. The fourth exercise stimulates executive functions. The user must arrange the products of a purchase into the rooms of a house.

Acceptance and Usability

A total of 19 people participated in the study (Table 1) and provided their opinions regarding the serious game, which are shown in Figure 6.
Table. Demographic characteristics of participants.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Value (N=19)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD)</td>
<td>75.3 (1.4)</td>
</tr>
<tr>
<td><strong>Gender, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Woman</td>
<td>16 (84)</td>
</tr>
<tr>
<td>Man</td>
<td>3 (16)</td>
</tr>
<tr>
<td><strong>Level of education, n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Secondary (high school)</td>
<td>19 (100)</td>
</tr>
<tr>
<td><strong>Smartphone experience (frequency of use), n (%)</strong></td>
<td></td>
</tr>
<tr>
<td>Once a week</td>
<td>19 (100)</td>
</tr>
</tbody>
</table>

Figure 6. Scores of older people who participated in the validation.
Principal Findings

From a cognitive point of view, the first effects of deterioration in older adults are directly related to the loss of the ability to carry out some instrumental ADL. However, there are multiple studies showing that people who read or play games are less likely to develop dementia or even Alzheimer disease [24,25]. In particular, the use of serious games has proven its value as a cognitive therapy for older people [26]. In the literature, different games have been designed and validated for cognitive stimulation [17,18], but it is not common for these games to use tasks focused on ADL. The aim of this work has been to describe how a serious game has been designed and implemented for cognitive stimulation of older adults by means of memory, naming, calculation, and organization exercises, which are key in ADL such as shopping. In this way, this serious game deals with the most frequent cognitive problems of older people indicated by health professionals.

According to the literature, it is not completely clear to what extent existing health care research explicitly distinguishes between gamification and serious games [27]. Serious games refer to the use of games and gaming technology for purposes other than just entertainment or fun, including health purposes. They can have direct or indirect positive physiological and psychological effects on people, which is precisely the objective of serious games in health and health care [28]. Our work applies to this definition because our serious game was designed for cognitive stimulation of older people with MCI and could have direct positive psychological effects on older people for the prevention or treatment of such cognitive impairment.

The preliminary pilot study carried out to validate the usability of the serious game showed that participants had a great opinion of this game and considered the theme interesting, entertaining, and useful for mental stimulation, so we can think a priori that the game could be well accepted among older people. Users found the serious game intuitive and aesthetically appealing; therefore, it meets the principles of simplicity and intuitiveness for the design of user interfaces for older people to avoid extracognitive load for the user [29].

Taking into account preferences of older people, game themes should meet their interest because they have a predilection for games related to real life [30]. In this sense, ecological validity was considered in the design of our serious game, since it is important for the validation of cognitive skills that influence functional tasks in real-world contexts [31]. Moreover, the fact that the difficulty of the game can be set by the professional to provide an achievable difficulty for each older adult user is important to motivate them to play and avoid frustration, anxiety, or negative emotions when playing [30].

Previous studies have described the benefits of using tablets for cognitive stimulation [32], but at the same time, rejection and barriers for older people in the use of these technologies have also been described [33]. In this regard, the results obtained in this study indicate that the game is easy to use and understand. However, these are preliminary findings as the study participants use smartphones in their daily live, which greatly reduces the rejection of this type of technology. Nevertheless, as future work, it is necessary to improve the interaction and facilitate the use of the game to avoid the rejection by older people, since results obtained from the questionnaires regarding the ease of use and instructions of the game were positive, but there was room for improvement.

In relation to the use of technologies by older adults and to digital health, the European Commission encourages improving the digitalization of health systems to fight health inequalities [34]. However, digital health services and devices are useless if consumers, in this case older people, do not have the skills or understanding to use them. For that, in the digital age, more than ever, literacy in digital health is a critical first step to improve the quality of life.

A feature of this serious game is that it provides feedback to the user at the end of each stage of the game, indicating whether the task was correctly completed or not and the type of mistake that was made. This agrees with Brox et al [35], who stated that when older people achieve their goal, feedback should be immediately provided. Moreover, to avoid frustration, which is another key aspect of the game, the system also provides encouraging messages during the completion of the game, so that the user can try again. This positive feedback for encouragement favors a successful experience of older user with the game [36], allowing them to achieve the goal with high motivation [30]. The use of narratives and the low complexity of the game are factors that motivate older people to play [37,38]. In this way, motivational mechanisms have been included in this serious game, which is the fundamental element for success of new training exercises.

According to the literature [39], game design features related to the game genre (GG), game nature (GN), and game development strategy (GDS) are necessary to develop a serious game. The GG covers sports, simulation, and strategy, among others. GN features include player perspective taking (first-person perspective vs third-person perspective), gameplay mode (multiplayer vs single player), type of scenery or in-game environment (realistic, fantasy themed, or simple), the presence or absence of playable characters, and the level of immersion applied to the use of immersive or nonimmersive virtual reality. GDS refers to custom-made serious games developed specifically for the study in question or direct-to-consumer approaches. In our work, regarding the GG, we can associate gender to the simulation of ADL, such as shopping, thus meeting this design criterion. In the case of GN, a first-person perspective, single-player experience, realistic scenario, the
absence of playable characters, and no immersion was implemented. According to these design features, those of interest that have not been included in our serious game and could be implemented and tested in the future works are as follows: (1) a multiplayer gameplay mode: in this way, older people could simulate going shopping together with their loved ones, family member, or friend, which is something they are probably used to in their life, thus increasing the feeling of sociability; and (2) changing the level of immersion to include the use of virtual reality: the use of virtual reality could be of interest, so that the older person is immersed in the game and can feel as if they are actually doing the shopping task, especially in the exercises that involve going to the supermarket, such as asking the shop assistant for the products, moving through the sections of the supermarket and filling the shopping basket with the products on the list, or even arriving home with the shopping basket and moving through the different sections of the house to place the products. All of these features would involve older people in a more realistic environment. However, the effort (human and material resources) involved in implementing these features, as well as the acceptance of older people to the use of this type of virtual reality technologies, either with glasses or simulating on a PC with a keyboard and mouse, would have to be carefully studied. Finally, regarding the third game design feature of GDS, our serious game has been implemented solely for our study, but it will be freely available for download and use by older people and health care institutions who wish to use it.

This serious game presents an alternative format to traditional interventions for older people, so we suggest defining a protocol to assess the effects on the cognitive function of older adults with a long-term intervention after a period of time to test the effectiveness of this serious game.

Limitations
The study has the limitation that the pilot study used a small sample size belonging to a limited geographical area, which may influence the generalizability of results. Therefore, it is necessary to extend these results with a larger sample size, including people of different cultural levels, with different experiences in the use of tactile devices, as well as people with different levels of cognitive impairment, to test the robustness of our findings.

Comparison With Prior Work
Ecological validity has been taken into account in the design of the serious game in this work, with tasks focused on ADL such as shopping. This is not common in the different games designed and validated for cognitive stimulation found in the literature, some of which focused on sport simulations, quizzes with text and images, music tools, arithmetic and reading calculations, etc [17], whereas others use a fantasy world setting, minigames, puzzles, or scenarios to cope with stressful or negative situations [18]. The innovative factor of including ecological scenarios of ADL tasks favors a greater interest and acceptance by older people and could help to assess the cognitive status to prevent or control the progress of the cognitive impairment.

Conclusions
In this work, an innovative serious game for cognitive stimulation of older people has been designed and developed, focusing on the ADL of shopping, incorporating motivational elements, and allowing for difficulty adaptability. The set of exercises included in the serious game have been described, including the theoretical basis on which each exercise has been implemented, which deal with the most frequent cognitive problems of older people associated with episodic declarative memory, naming, calculation, and organization. This study could serve as a basis for future serious games for cognitive stimulation of older people that may benefit from the knowledge obtained about the design strategies followed.

A pilot usability study carried out with older adults has shown that this serious game is intuitive, provides a high level of entertainment, and is useful for its application in daily life. In conclusion, according to the preliminary results obtained, we think that the serious game could be widely accepted by older people and, therefore, could become a tool that contributes to delaying their deterioration and increasing their independence. In this way, the serious game developed could contribute to increase the low degree of awareness of older people regarding the importance of leading an active life.

Acknowledgments
The authors would like to gratefully acknowledge all the people from Spain and Portugal who willingly participated in the study. Moreover, they would like to thank the Pixel Ratio SL company for the technical development of the serious game. This study was carried out under the EuroAGE+ Project (0124_EUROAGE_MAS_4_E), “Red Internacional de Investigación, Innovación y Transferencia de Tecnologías para la promoción del envejecimiento activo,” financed jointly by the European Regional Development Fund (ERDF) through the Operational Program of Cross-border Cooperation Spain - Portugal (POCTEP) 2021-2027. The paper was cofinanced by the Spanish Ministry of Science and Innovation TED2021-131739-C22, supported by MCIN/AEI/10.13039/501100011033 and the European Union “NextGenerationEU”/PRTR.

Conflicts of Interest
None declared.

References


12. Blasco SV, Moreno MJC. Las nuevas tecnologías en el tratamiento de enfermos con demencias y la conveniencia de incluir estos contenidos en los estudios de disciplinas de ciencias de la salud [Article in Spanish], Psiquiatría 2021;25 [FREE Full text]


Abbreviations

ADL: activities of daily living
GDS: game development strategy
GG: game genre
GN: game nature
ICT: information and communication technology
MCI: mild cognitive impairment

Edited by R Yang; submitted 26.07.22; peer-reviewed by A Sheikhtaheri, B Chaudhry, D Chrimes, K Nalin, M Jordan-Marsh, SF Fung, T Baranowski; revised version received 19.10.23; accepted 26.10.23; published 04.04.24.

Please cite as:
Ortega Morán JF, Pagador JB, Gilete Preciado V, Moyano-Cuevas JL, Rodríguez Domínguez T, Santurino Muñoz M, Sánchez Margallo FM
A Serious Game for Cognitive Stimulation of Older People With Mild Cognitive Impairment: Design and Pilot Usability Study
JMIR Aging 2024;7:e41437
URL: https://aging.jmir.org/2024/1/e41437
doi: 10.2196/41437
Association Between Sleep Efficiency Variability and Cognition Among Older Adults: Cross-Sectional Accelerometer Study

Collin Sakal¹, BA, MSc; Tingyou Li¹; Juan Li², BA, MA, PhD; Can Yang³, BS, MPhil, PhD; Xinyue Li¹, BA, MS, PhD

¹ ² ³

Corresponding Author:
Xinyue Li, BA, MS, PhD

Abstract

Background: Sleep efficiency is often used as a measure of sleep quality. Getting sufficiently high-quality sleep has been associated with better cognitive function among older adults; however, the relationship between day-to-day sleep quality variability and cognition has not been well-established.

Objective: We aimed to determine the relationship between day-to-day sleep efficiency variability and cognitive function among older adults, using accelerometer data and 3 cognitive tests.

Methods: We included older adults aged >65 years with at least 5 days of accelerometer wear time from the National Health and Nutrition Examination Survey (NHANES) who completed the Digit Symbol Substitution Test (DSST), the Consortium to Establish a Registry for Alzheimer’s Disease Word-Learning subtest (CERAD-WL), and the Animal Fluency Test (AFT). Sleep efficiency was derived using a data-driven machine learning algorithm. We examined associations between sleep efficiency variability and scores on each cognitive test adjusted for age, sex, education, household income, marital status, depressive symptoms, diabetes, smoking habits, alcohol consumption, arthritis, heart disease, prior heart attack, prior stroke, activities of daily living, and instrumental activities of daily living. Associations between average sleep efficiency and each cognitive test score were further examined for comparison purposes.

Results: A total of 1074 older adults from the NHANES were included in this study. Older adults with low average sleep efficiency exhibited higher levels of sleep efficiency variability (Pearson \( r = -0.63 \)). After adjusting for confounding factors, greater average sleep efficiency was associated with higher scores on the DSST (per 10% increase, \( \beta = 2.25, 95\% \text{ CI } 0.61 \text{ to } 3.90 \)) and AFT (per 10% increase, \( \beta = 0.91, 95\% \text{ CI } 0.27 \text{ to } 1.56 \)). Greater sleep efficiency variability was univariably associated with worse cognitive function based on the DSST (per 10% increase, \( \beta = -3.34, 95\% \text{ CI } -5.33 \text{ to } -1.34 \)), CERAD-WL (per 10% increase, \( \beta = -1.00, 95\% \text{ CI } -1.79 \text{ to } -0.21 \)), and AFT (per 10% increase, \( \beta = -1.02, 95\% \text{ CI } -1.68 \text{ to } -0.36 \)). In fully adjusted models, greater sleep efficiency variability remained associated with lower DSST (per 10% increase, \( \beta = -2.01, 95\% \text{ CI } -3.62 \text{ to } -0.40 \)) and AFT (per 10% increase, \( \beta = -0.84, 95\% \text{ CI } -1.47 \text{ to } -0.21 \)) scores but not CERAD-WL (per 10% increase, \( \beta = -0.65, 95\% \text{ CI } -1.39 \text{ to } 0.08 \)) scores.

Conclusions: Targeting consistency in sleep quality may be useful for interventions seeking to preserve cognitive function among older adults.

(JMIR Aging 2024;7:e54353) doi:10.2196/54353

KEYWORDS

aging; cognition; accelerometer; sleep; sleep efficiency; geriatrics; gerontology; actigraphy; digital health; mhealth; cognitive impairments; mobile health; efficiency; variability; older adult; older adults; elder; elderly; older person; sleep quality; machine learning; quality of sleep; sleep

Introduction

Healthy sleep habits protect memory and cognitive function [1,2]. Sleep quality deteriorates with age, but older adults with cognitive impairments have worse sleep quality than their counterparts without impairments [3-5]. Lower sleep efficiency, a proxy for sleep quality, is associated with worse cognition among older adults [6]. The importance of getting sufficiently high-quality sleep to reduce individual risk of cognitive impairments has been reported [3,7]; however, the relationship between consistent sleep quality and cognition remains understudied. Because it is unreasonable to assume that older adults strictly adhere to a consistent sleep schedule on a nightly
basis, the relationship between day-to-day sleep efficiency variability and cognition must be examined.

This cross-sectional accelerometer study aimed to quantify associations between sleep efficiency variability and performance on 3 cognitive tests assessing memory, categorical verbal fluency, and sustained attention while adjusting for demographic factors, chronic diseases, smoking habits, alcohol consumption, cardiovascular risk factors, depressive symptoms, and measures of activities of daily living (ADL) and instrumental activities of daily living (IADL). We additionally fit models using average sleep efficiency metrics to compare any observed relationships between sleep efficiency variability and cognition to those between average sleep efficiency and cognition.

Methods

Data Source and Study Design

Data from the US National Health and Nutrition Examination Survey (NHANES) 2011-2014 waves [8] were used, during which a subset of participants wore an ActiGraph GT3X+ device that objectively measured activity levels over 7 consecutive days immediately after all nonaccelerometer data were collected. Participants aged >60 years were also administered cognitive tests during the 2011-2014 waves. We excluded participants aged <65 years, without complete cognitive test data, or without at least 5 days of accelerometer wear time.

Ethical Considerations

All NHANES participants provided informed consent, and ethics approval was granted by the National Center for Health Statistics Research Ethics Review Board (protocol #2011-17).

Measuring Cognition

The NHANES 2011-2014 waves include 3 cognitive tests: the Digit Symbol Substitution Test (DSST), the Consortium to Establish a Registry for Alzheimer’s Disease Word-Learning subtest (CERAD-WL), and the Animal Fluency Test (AFT; Table 1) [9]. For the AFT, CERAD-WL, and DSST, higher scores correspond to better cognition.

Table 1. The 3 cognitive tests included in the 2011-2014 waves of the US National Health and Nutrition Examination Survey.

<table>
<thead>
<tr>
<th>Cognitive test</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DSST&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Tests processing speed, sustained attention, and working memory. Scores range from 0 to 133 where 1 point is awarded for each symbol correctly written below its corresponding number based on a key mapping the symbols to the numbers.</td>
</tr>
<tr>
<td>CERAD-WL&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Measures immediate and delayed word recall. Three rounds of immediate recall and 1 round of delayed recall using lists of 10 unrelated words comprise the CERAD-WL. Scores on the CERAD-WL correspond to the number of correctly recalled words across all 3 rounds.</td>
</tr>
<tr>
<td>AFT&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Measures verbal categorical fluency and requires participants to name as many animals as possible in a 1-minute period.</td>
</tr>
</tbody>
</table>

<sup>a</sup>DSST: Digit Symbol Substitution Test.

<sup>b</sup>CERAD-WL: Consortium to Establish a Registry for Alzheimer’s Disease Word-Learning subtest.

<sup>c</sup>AFT: Animal Fluency Test.

Deriving Sleep Metrics

Sleep efficiency—a proxy for sleep quality—is the ratio of time asleep to the total time between sleep onset and final sleep offset; possible values range 0-1 with higher values corresponding to better-quality sleep. Nightly sleep efficiency values were derived using an unsupervised hidden Markov model that identifies sleep-wake states using a data-driven machine learning approach [10]. The hidden Markov model algorithm has been validated against gold-standard polysomnography with 85.7% accuracy, 99.3% sensitivity, and better performance than commonly used supervised algorithms [10]. Sleep efficiency variability was defined as the SD of sleep efficiency across all nights of valid accelerometer data. For comparison purposes, we further derived each participant’s average sleep efficiency.

Additional Covariates

To account for potential confounders, we gathered each participant’s age, sex, education, marital status, household income, smoking habits, current alcohol consumption, depressive symptoms, measures of functional independence, history of heart attack, history of stroke, and diagnoses of arthritis, heart disease, and diabetes. Depressive symptoms were quantified using Patient Health Questionnaire-9 scores [11]. A functional independence score was derived by summing responses to 20 ADL and IADL questions. Participants were categorized as current, former, or never smokers and drinkers. An explicit explanation of how each covariate was defined can be found in Multimedia Appendix 1. Participants with missing data were excluded to enable a complete-case analysis.

Statistical Analysis

Participant characteristics were reported using the means and SDs for numeric variables and counts and percentages for categorical variables. We first examined the relationship between mean and day-to-day sleep efficiency variability using the Pearson r correlation coefficient and a scatterplot. Thereafter, using cutoffs from previous studies [12], we plotted the distribution of sleep efficiency variability stratified by normal versus low (≥0.85 vs <0.85) sleep efficiency.
We first examined univariable associations between sleep efficiency variability and DSST, CERAD-WL, and AFT scores. Demographic models were adjusted for age, sex, education, marital status, and household income. Finally, the full models in this study were further adjusted for depressive symptoms, ADL and IADL scores, smoking habits, alcohol consumption, diabetes, arthritis, heart disease, history of stroke, and history of heart attack. All univariable, demographic, and full models were refit using average sleep efficiency instead of day-to-day variability for comparison purposes. A sensitivity analysis was then conducted where we excluded extreme outliers (observations ≤1st or ≥99th quantile) for both average sleep efficiency and day-to-day variability. Models with both average and sleep efficiency variability were also examined (Multimedia Appendix 2).

**Results**

**Descriptive Statistics**

In total, 1074 NHANES participants were included (mean age 72.3, SD 5.2 years; females: n=528, 49%), among whom 97.8% (n=1051) had ≥7 nights of sleep data (Table 2 and Table S1 and Figure S1 in Multimedia Appendix 1). The average sleep efficiency in the cohort was 0.94 (SD 0.05), while the average DSST, CERAD-WL, and AFT scores were 46.7 (SD 16.0), 25.0 (SD 6.29), and 16.8 (SD 5.25), respectively. The correlation between mean and day-to-day sleep efficiency variability was −0.63 (Figure 1). We found that older adults with low average sleep efficiency had higher levels of sleep efficiency variability than those with normal sleep efficiency levels (Figure 2).

**Table**. Demographic, sleep, and cognitive characteristics of older adults (N=1074) with valid accelerometer and cognitive test data who were part of the National Health and Nutrition Examination Survey.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nights of sleep data, mean (SD)</td>
<td>7.8 (0.47)</td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>72.3 (5.2)</td>
</tr>
<tr>
<td>Sex, n (%), Male</td>
<td>546 (0.51)</td>
</tr>
<tr>
<td>Sex, n (%), Female</td>
<td>528 (0.49)</td>
</tr>
<tr>
<td>Education, n (%), Less than ninth grade</td>
<td>95 (0.09)</td>
</tr>
<tr>
<td>Education, n (%), Some high school</td>
<td>141 (0.13)</td>
</tr>
<tr>
<td>Education, n (%), High school graduate or GEDa</td>
<td>245 (0.23)</td>
</tr>
<tr>
<td>Education, n (%), Some college or associate’s degree</td>
<td>307 (0.29)</td>
</tr>
<tr>
<td>Education, n (%), College graduate or above</td>
<td>286 (0.27)</td>
</tr>
<tr>
<td>Marital status, n (%), Married</td>
<td>613 (0.57)</td>
</tr>
<tr>
<td>Marital status, n (%), Widowed</td>
<td>230 (0.21)</td>
</tr>
<tr>
<td>Marital status, n (%), Divorced</td>
<td>145 (0.14)</td>
</tr>
<tr>
<td>Marital status, n (%), Separated</td>
<td>19 (0.02)</td>
</tr>
<tr>
<td>Marital status, n (%), Never married</td>
<td>42 (0.04)</td>
</tr>
<tr>
<td>Marital status, n (%), Living with partner</td>
<td>25 (0.02)</td>
</tr>
<tr>
<td>Sleep efficiency variability, mean (SD)</td>
<td>0.04 (0.05)</td>
</tr>
<tr>
<td>Average sleep efficiency, mean (SD)</td>
<td>0.94 (0.05)</td>
</tr>
<tr>
<td>DSSTb score, mean (SD)</td>
<td>46.7 (16.0)</td>
</tr>
<tr>
<td>CERAD-WLc score, mean (SD)</td>
<td>25.0 (6.29)</td>
</tr>
<tr>
<td>AFTd score, mean (SD)</td>
<td>16.8 (5.25)</td>
</tr>
</tbody>
</table>

aGED: General Educational Development.
bDSST: Digit Symbol Substitution Test.
cCERAD-WL: Consortium to Establish a Registry for Alzheimer’s Disease Word-Learning subtest.
dAFT: Animal Fluency Test.
Figure 1. Scatterplot with a fitted regression line of average versus day-to-day variability for sleep efficiency.
Associations

In the univariable models, greater sleep efficiency variability was associated with lower scores on the DSST (per 10% increase, $\beta = -3.34$, 95% CI $-5.33$ to $-1.34$), CERAD-WL (per 10% increase, $\beta = -1.00$, 95% CI $-1.79$ to $-0.21$), and AFT (per 10% increase, $\beta = -1.02$, 95% CI $-1.68$ to $-0.36$). In the full models, greater sleep efficiency variability was associated with lower scores on the DSST (per 10% increase, $\beta = -2.01$, 95% CI $-3.62$ to $-0.40$) and AFT (per 10% increase, $\beta = -0.84$, 95% CI $-1.47$ to $-0.21$) but not the CERAD-WL (per 10% increase, $\beta = -0.65$, 95% CI $-1.39$ to 0.08; Tables 3-5). Conversely, greater average sleep efficiency was associated with higher scores on the DSST (per 10% increase, $\beta = 2.25$, 95% CI 0.61-3.90) and AFT (per 10% increase, $\beta = 0.91$, 95% CI 0.27-1.56) but not the CERAD-WL (per 10% increase, $\beta = 0.46$, 95% CI $-0.29$ to 1.21) in the full models. In the sensitivity analysis, after excluding extreme averages and sleep efficiency variability outliers, all significant results observed in the full models remained significant (Multimedia Appendix 3).
Table. Associations between day-to-day variability and average sleep efficiency with Digit Symbol Substitution Test (DSST) scores.

<table>
<thead>
<tr>
<th>Model covariates</th>
<th>Association with DSST scores</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β (95% CI)(^a)</td>
<td></td>
</tr>
<tr>
<td>Sleep efficiency variability</td>
<td>−3.34 (−5.33 to −1.34)</td>
<td>.001</td>
</tr>
<tr>
<td>Mean sleep efficiency</td>
<td>4.28 (2.27 to 6.28)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Demographics + sleep efficiency variability</td>
<td>−2.04 (−3.69 to −0.39)</td>
<td>.02</td>
</tr>
<tr>
<td>Demographics + mean sleep efficiency</td>
<td>2.65 (0.97 to 4.32)</td>
<td>.002</td>
</tr>
<tr>
<td>Full model + sleep efficiency variability</td>
<td>−2.01 (−3.62 to −0.40)</td>
<td>.02</td>
</tr>
<tr>
<td>Full model + average sleep efficiency</td>
<td>2.25 (0.61 to 3.90)</td>
<td>.007</td>
</tr>
</tbody>
</table>

\(^a\)Coefficients are reported per 10% increase.

Table. Associations of day-to-day variability and average sleep efficiency with Consortium to Establish a Registry for Alzheimer’s Disease Word-Learning subtest (CERAD-WL) scores.

<table>
<thead>
<tr>
<th>Model covariates</th>
<th>Association with CERAD-WL scores</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β (95% CI)(^a)</td>
<td></td>
</tr>
<tr>
<td>Sleep efficiency variability</td>
<td>−1.00 (−1.79 to −0.21)</td>
<td>.01</td>
</tr>
<tr>
<td>Mean sleep efficiency</td>
<td>0.85 (0.06 to 1.65)</td>
<td>.04</td>
</tr>
<tr>
<td>Demographics + sleep efficiency variability</td>
<td>−0.70 (−1.43 to 0.03)</td>
<td>.06</td>
</tr>
<tr>
<td>Demographics + mean sleep efficiency</td>
<td>0.52 (−0.23 to 1.26)</td>
<td>.18</td>
</tr>
<tr>
<td>Full model + sleep efficiency variability</td>
<td>−0.65 (−1.39 to 0.08)</td>
<td>.08</td>
</tr>
<tr>
<td>Full model + average sleep efficiency</td>
<td>0.46 (−0.29 to 1.21)</td>
<td>.23</td>
</tr>
</tbody>
</table>

\(^a\)Coefficients are reported per 10% increase.

Table. Associations of day-to-day variability and average sleep efficiency with Animal Fluency Test (AFT) scores.

<table>
<thead>
<tr>
<th>Model covariates</th>
<th>Association with AFT scores</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β (95% CI)(^a)</td>
<td></td>
</tr>
<tr>
<td>Sleep efficiency variability</td>
<td>−1.02 (−1.68 to −0.36)</td>
<td>.002</td>
</tr>
<tr>
<td>Mean sleep efficiency</td>
<td>1.08 (0.42 to 1.74)</td>
<td>.001</td>
</tr>
<tr>
<td>Demographics + sleep efficiency variability</td>
<td>−0.85 (−1.48 to −0.22)</td>
<td>.009</td>
</tr>
<tr>
<td>Demographics + mean sleep efficiency</td>
<td>1.02 (0.38 to 1.66)</td>
<td>.002</td>
</tr>
<tr>
<td>Full model + sleep efficiency variability</td>
<td>−0.84 (−1.47 to −0.21)</td>
<td>.009</td>
</tr>
<tr>
<td>Full model + average sleep efficiency</td>
<td>0.91 (0.27 to 1.56)</td>
<td>.005</td>
</tr>
</tbody>
</table>

\(^a\)Coefficients are reported per 10% increase.

**Discussion**

**Principal Results and Comparisons With Prior Work**

This study shows that older adults with higher sleep efficiency variability scored worse on the DSST and AFT after adjusting for demographic factors, chronic diseases, smoking habits, alcohol consumption, depressive symptoms, cardiovascular risk factors, and ADL and IADL scores. Effect sizes for average and sleep efficiency variability were similar in magnitude but in opposite directions, with greater variability being associated with lower DSST and AFT scores, while greater average sleep efficiency was associated with higher scores.

A previous accelerometer study found that greater sleep efficiency variability was associated with lower scores on serial subtraction tests and memory questionnaires [13]. However, the study was limited by a small sample (n<50) and did not consider relevant confounders such as chronic diseases, ADL and IADL, smoking habits, and alcohol consumption. Another study found that greater sleep efficiency variability was associated with a greater β-amyloid burden—a biomarker for Alzheimer disease—but was again limited by a small sample (n=52) [14]. Compared to existing studies, our work, using a larger cohort accounting for more confounders, provides evidence that greater sleep efficiency variability is associated with worse cognitive function among older adults. Furthermore, we found that effect sizes for sleep efficiency variability and average sleep efficiency were similar but in opposite directions,
suggesting that getting sufficient and consistent high-quality sleep may be useful targets for interventions seeking to preserve cognitive function among older adults.

**Limitations**

Given the cross-sectional nature of this study, we could not examine causal relationships. Bidirectional associations exist between certain forms of cognitive impairment and sleep disturbances [15], but they could not be examined herein. Future studies with longitudinal designs are therefore needed to further examine whether high sleep efficiency variability causally influences cognitive function or vice versa. Future studies may also benefit from collecting accelerometer data over longer durations, which more reliably measures sleep parameters [16], and from analyzing data across different countries to assess the generalizability of the findings presented herein. Lastly, polysomnography, the gold standard for measuring sleep parameters, was not used in this study. However, the use of wrist-worn accelerometers helped us assess sleep under natural living conditions, which intrusive polysomnography does not permit.

**Conclusions**

Greater day-to-day sleep efficiency variability was associated with lower scores on 2 cognitive tests in this study. Our work may motivate future causal inference studies seeking to determine if consistency in sleep quality is a viable target for preserving cognitive function among older adults.

**Acknowledgments**

This work was supported by the City University of Hong Kong, Hong Kong SAR, China, with internal research grants (#9610473 and #7005892). The funder played no role in study design, data collection, analysis, and interpretation of data or writing of this manuscript.

**Authors' Contributions**

CS, XL, and JL designed the study. CS and TL performed the statistical analyses with supervision from XL and JL. The manuscript was primarily written by CS with oversight from XL, CY, and JL.

**Conflicts of Interest**

None declared.

**Multimedia Appendix 1**

Participant flowchart, cohort characteristics, covariate definitions.

[DOCX File, 106 KB - aging_v7i1e54353_app1.docx ]

**Multimedia Appendix 2**

Additional models including both average and sleep efficiency variability

[DOCX File, 26 KB - aging_v7i1e54353_app2.docx ]

**Multimedia Appendix 3**

Sensitivity analysis.

[DOCX File, 26 KB - aging_v7i1e54353_app3.docx ]

**References**


Abbreviations
ADL: activities of daily living
AFT: Animal Fluency Test
CERAD-WL: Consortium to Establish a Registry for Alzheimer’s Disease Word-Learning subtest
DSST: Digit Symbol Substitution Test
IADL: instrumental activities of daily living
NHANES: National Health and Nutrition Examination Survey
The Implementation Success of Technology-Based Counseling in Dementia Care: Scoping Review

Dorothee Bauernschmidt1*, Dr rer medic; Janina Wittmann1*, MSc; Julian Hirt2,3, Dr rer medic; Gabriele Meyer1, Dr phil; Anja Bieber1, Dr rer medic

1Institute of Health and Nursing Science, University Medicine Halle, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany
2Center for Dementia Care, Institute of Nursing Science, Department of Health, Eastern Switzerland University of Applied Sciences, St Gallen, Switzerland
3Department of Clinical Research, University Hospital Basel, University of Basel, Basel, Switzerland
*these authors contributed equally

Corresponding Author:
Anja Bieber, Dr rer medic
Institute of Health and Nursing Science
University Medicine Halle
Martin Luther University Halle-Wittenberg
Magdeburger Straße 8
Halle (Saale), 06112
Germany
Phone: 49 345 557 4427
Email: anja.bieber@uk-halle.de

Abstract

Background: Implementing technology-based counseling as a complex intervention in dementia care poses challenges such as adaptation to stakeholders’ needs and limited resources. While studies have examined the effectiveness of technology-based counseling, its successful implementation remains largely unexplored.

Objective: We aimed to review the knowledge about the implementation success of technology-based counseling interventions for people with dementia and their informal caregivers.

Methods: We conducted a scoping review and systematically searched CINAHL, the Cochrane Library including the Cochrane Central Register of Controlled Trials, MEDLINE, PsycINFO, and Web of Science Core Collection databases (April 2021) in combination with citation searching and web searching (November 2021). Studies reporting on technology-based counseling interventions for people with dementia or their informal caregivers were included, irrespective of the design. We used the conceptual framework for implementation outcomes to operationalize implementation success and applied the outcomes acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, penetration, and sustainability as categories to inform data extraction. We identified dimensions within the categories and synthesized results narratively and graphically.

Results: We included 52 publications reporting on 27 technology-based counseling interventions. The studies were conducted in 9 countries and published between 1993 and 2021. As the design of the included studies varied, the number of participants and the type of data reported varied as well. The intervention programs were heterogeneous and ranged from single counseling interventions (such as helpline services) to counseling as part of a multicomponent program. Telephone, email, videoconferencing, social media (respectively chats), and web-based platforms were used for delivering counseling. We found data on appropriateness for all interventions and data on acceptability for most interventions, describing aspects such as consumer-perceived usefulness and helpfulness of services, as well as satisfaction. Information on the other categories of adoption, feasibility, fidelity, implementation cost, penetration, and sustainability was fragmented.

Conclusions: The scope and depth of information on conceptual categories of the implementation success of technology-based counseling for people with dementia and informal caregivers varied. The data only partially covered the concept of implementation success, which highlights the need for a systematic evaluation accompanying the implementation. The application of theoretical approaches for implementation and adherence to the framework for developing and evaluating complex interventions are required to promote the implementation of complex interventions and to comprehensively assess implementation success.

Trial Registration: PROSPERO CRD42021245473; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=245473
Another challenge in the implementation of complex capabilities to support innovation or preferences for classically In the wider context, aspects such as stakeholders' limited time constraints can also act as barriers to implementation [16]. Furthermore, financial and the lack of staff resources or the lack of interaction with staff information is disclosed [16]. In implementing organizations, users' concerns about security issues, especially when personal counseling offers the possibility of providing support without the risk of infection [10,14]. Counseling for people with dementia and their informal caregivers is a supportive intervention that can be defined as conversational therapy in which a trained therapist listens to the person, enhances the individual’s ability to cope with the effects of dementia, and provides information and education [7,8]. Counselors provide information about the disease and support service options, offer the opportunity to share and discuss feelings or problems, and can enhance problem-solving and coping skills [9,10]. In the context of dementia, counseling services provided by professionals and tailored to individuals aim at various outcomes, such as reducing depressive symptoms and perceived burden, improving quality of life and self-efficacy, and encouraging the uptake of support services in the community [11]. Counseling interventions are therefore complex interventions comprising several components, such as specific training for providers, tools and instruments to individualize services, or different ways to access the available services [9,11]. Information and communication technologies may improve accessibility to counseling [12,13], and this aspect gained importance during the COVID-19 pandemic when remote counseling offered the possibility of providing support without the risk of infection [10,14].

To successfully translate complex interventions into practice, various challenges need to be overcome. Implementation can be defined as an “actively planned and deliberately initiated effort with the intention to bring a given intervention into policy and practice within a particular setting” [15]. The challenges described in implementing eHealth interventions in dementia care include adapting the interventions to match the skills and abilities of the target population, achieving user friendliness in the context of rapidly evolving technology, and addressing users’ concerns about security issues, especially when personal information is disclosed [16]. In implementing organizations, the lack of staff resources or the lack of interaction with staff as well as the reluctance of the provider to use the technology may impede implementation [16]. Furthermore, financial and time constraints can also act as barriers to implementation [16]. In the wider context, aspects such as stakeholders’ limited capabilities to support innovation or preferences for classically delivered care were identified as significant barriers [16].

Another challenge in the implementation of complex interventions is to operationalize and measure the success or effectiveness of implementation efforts. In a Cochrane Review on remotely delivered information, training, and support (including counseling) for informal caregivers of people with dementia, the authors found information on various aspects indicating implementation success such as acceptability, user satisfaction, or fidelity of implementation. As the indicators and assessment areas varied substantially across studies, data synthesis could not be conducted [9]. There are theoretical approaches that facilitate a comprehensive evaluation of implementation efforts [15]. One of them is the conceptual framework for implementation outcomes introduced by Proctor et al [17], which provides implementation-specific outcomes for assessing implementation success.

While there are studies investigating the effectiveness of technology-based counseling [8,9,11], the extent to which these interventions are successfully implemented remains unexplored. To address this gap in knowledge, we aimed to review the evidence and pursue the question of what is known about the implementation success of technology-based counseling interventions for people with dementia or their informal caregivers.

Methods

Design

According to the methodological approach of scoping reviews [18], we aimed at mapping evidence of implementation success to provide a comprehensive overview. We followed the methodological guidance [18] of the Joanna Briggs Institute and structured our report according to the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) [19].

Protocol and Registration

This scoping review is embedded in the project “Technology-based counselling in dementia (TeCoDem),” for which a protocol has been developed [7] and registered with the international prospective register of systematic reviews (PROSPERO CRD42021245473; see the section Deviations From the Protocol).

Eligibility Criteria

We included studies, irrespective of their design in English and German, that reported on technology-based counseling interventions for people with any type and severity of dementia or their informal caregivers. Interventions had to be tailored to individuals and provided remotely by professionals using various information technologies. Studies on people with mild cognitive impairment as well as studies on standardized interventions, such as cognitive behavioral therapy, on genetic counseling, and on counseling regarding diagnostics or screening for dementia were excluded. We also excluded studies describing interventions that focus mainly on care coordination or case
management. Furthermore, publications reporting exclusively on the development of interventions were excluded [7].

**Information Sources and Search Strategy**

We searched CINAHL, the Cochrane Library including the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE via PubMed, PsycINFO via Ovid, and the Web of Science Core Collection databases (last search: April 22, 2021) without any filters and limiters. We systematically developed a search strategy that contained 3 components: dementia, technology, and counseling. Corresponding search terms and synonyms (eg, dementia/Alzheimer, technology/electronic, counselling/counseling/consultation) were identified through an orienting search using MEDLINE via PubMed, and we checked entry terms given in the Medical Subject Headings browser. The strategy was peer-reviewed by applying the Peer Review of Electronic Search Strategies [20]. In addition, we performed forward and backward citation searches of included studies and pertinent reviews via Scopus (last search: October 7, 2021) and a web search via Google and Google Scholar (last search: November 26, 2021) [21,22]. Full database-specific search strategies are provided elsewhere [7].

**Selection of Sources of Evidence**

Titles, abstracts, and full texts were independently screened by 2 reviewers (out of AB, JH, and DB) using the Rayyan web application (Rayyan) [23]. Any discrepancies in the decisions were resolved by discussions within the review team.

### Data Charting Process and Data Items

A targeted and uniform extraction sheet was developed and consented to by the research team. We extracted study and design characteristics (year of publication, country where the study was conducted, objectives, number of participants or contacts) and assessed the technology-based counseling interventions by applying criteria from the Template for Intervention Description and Replication checklist [24] and from the revised Criteria for Reporting the Development and Evaluation of Complex Interventions guideline [25] to obtain a comprehensive overview of the interventions and their components. Data extraction on implementation success was guided by the conceptual framework for implementation outcomes, which comprises the implementation outcomes acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, penetration, and sustainability [17]. We adapted the definitions of the outcomes by specifying the intervention of interest and adjusting it to the research interest of our scoping review. Adaptations were consented to by the review team. The original and adapted definitions are listed in **Table 1**.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Original definition according to Proctor et al [17]</th>
<th>Adapted definition for our review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td>“Acceptability is the perception among implementation stakeholders that a given treatment, service, practice, or innovation is agreeable, palatable, or satisfactory.”</td>
<td>Acceptability is the perception among implementation stakeholders of technology-based counseling that the intervention is agreeable, palatable, or satisfactory.</td>
</tr>
<tr>
<td>Adoption</td>
<td>“Adoption is defined as the intention, initial decision, or action to try or employ an innovation or evidence-based practice.”</td>
<td>Adoption is defined as the intention, initial decision, or action to try or employ a technology-based counseling intervention.</td>
</tr>
<tr>
<td>Appropriateness</td>
<td>“Appropriateness is the perceived fit, relevance, or compatibility of the innovation or evidence based practice for a given practice setting, provider, or consumer; and/or perceived fit of the innovation to address a particular issue or problem.”</td>
<td>Appropriateness is the perceived fit, relevance, or compatibility of the technology-based counseling intervention for the given practice setting, provider, and consumer; and/or perceived fit of the intervention to address a particular issue or problem.</td>
</tr>
<tr>
<td>Feasibility</td>
<td>“Feasibility is defined as the extent to which a new treatment, or an innovation, can be successfully used or carried out within a given agency or setting (Karsh 2004).”</td>
<td>Feasibility is defined as the extent to which the technology-based counselling intervention can be successfully used or carried out within a given agency or setting.</td>
</tr>
<tr>
<td>Fidelity</td>
<td>“Fidelity is defined as the degree to which an intervention was implemented as it was prescribed in the original protocol or as it was intended by the program developers (Dusenbury et al. 2003; Rabin et al. 2008).”</td>
<td>Fidelity is the degree to which a technology-based counseling intervention was implemented as it was prescribed in the original protocol or as it was intended by the program developers.</td>
</tr>
<tr>
<td>Implementation cost</td>
<td>“Cost (incremental or implementation cost) is defined as the cost impact of an implementation effort.”</td>
<td>Implementation cost is the cost impact of an implementation effort.</td>
</tr>
<tr>
<td>Penetration</td>
<td>“Penetration is defined as the integration of a practice within a service setting and its subsystems.”</td>
<td>Penetration is defined as the integration of a technology-based counseling intervention within a service setting.</td>
</tr>
<tr>
<td>Sustainability</td>
<td>“Sustainability is defined as the extent to which a newly implemented treatment is maintained or institutionalized within a service setting’s ongoing, stable operations.”</td>
<td>Sustainability is defined as the extent to which an implemented technology-based counseling intervention is maintained or institutionalized within an organization’s ongoing, stable operations.</td>
</tr>
</tbody>
</table>

**Table 1.** Original and adapted definitions of implementation outcomes.

Quotations from the included studies were extracted and assigned to the outcomes by 1 reviewer (out of DB or JW) and cross-checked for accuracy by another reviewer (out of DB or JW). Any discrepancies were resolved by consensus between these 2 reviewers.

**Synthesis of Results**

Data on the implementation success of technology-based counseling interventions were summarized by applying the framework mentioned in the preceding section [17]. The following 8 conceptually distinct implementation outcomes...
were used as conceptual categories to operationalize implementation success: acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, penetration, and sustainability.

Our approach to data synthesis involved the following steps:

1. Concept specification: identification of dimensions in the definition of each conceptual category: as these conceptual categories incorporate various aspects [17] and therefore represent multidimensional concepts, a concept specification was performed by determining the dimensions of the categories [26,27]. Dimensions are defined as characteristics according to which empirical facts can be distinguished [28]. The 2 reviewers (DB and JW) independently identified dimensions matching the attributes of the conceptual categories [27] described by Proctor et al [17] with the characteristics of the extracted data. Consensus on the dimensions was reached through discussion between the 2 reviewers.

2. Reduction of data and assignment to dimensions in analysis matrices: the extracted data were reduced without paraphrasing and assigned to the dimensions using tables as analysis matrices.

3. Specification of the level of analysis: we specified the level at which data were provided (level of analysis: consumer, provider or providing institution, organization, setting, and administration), as indicated by Proctor et al [17].

4. Narratively synthesizing of findings and graphical presentation: findings were narratively synthesized and presented in the form of a net diagram.

Each synthesis step was cross-checked (DB and JW) and consent was obtained from the review team.

Study characteristics and characteristics of the included interventions are presented in narrative and tabular forms.

**Deviations From the Protocol**

The prespecified method of conducting a Qualitative Comparative Analysis on the conditions of successful implementation of technology-based counseling interventions [7] could not be realized because of the heterogeneity of the data found in the literature. In addition, reports that were not written in English or German were excluded because of a lack of professional translation resources.

**Results**

**Selection of Sources of Evidence**

The electronic database yielded 6387 records. After removing duplicates, we screened the titles and the abstracts of 3775 records, reviewed 277 full texts for eligibility, and included 35 records. We identified 3614 records from additional sources and assessed 151 full texts, of which 17 were included. Finally, 52 publications [29-80] reporting on 27 technology-based counseling interventions were included (Figure 1).

**Characteristics of Sources of Evidence**

Interventions ranged from single counseling interventions, such as helpline services, to multicomponent programs combining nontechnology-based components, such as day care for people with dementia, with technology-based counseling for informal caregivers. All 27 interventions [29-80] focused on informal caregivers, mostly in combination with people with dementia, and 4 also addressed professional caregivers [29-32,34,35,37,40]. Counseling was provided by professionals...
from different disciplines, such as psychologists, geriatricians, or nurses. If volunteers were involved, they were professionally supervised or had received training. The technologies used for delivery included telephone, email, videoconferencing, social media, and chats, as well as web-based platforms. Additional (personalized) information material was frequently offered and delivered via email or postal mail. We differentiated the following types of interventions (refer to Table S1 in Multimedia Appendix 1): counseling via telephone or email and counseling via videoconferencing; web-based psychosocial intervention: information, communication, and counseling; videoconference- or telephone-based counseling combined with tele-monitoring or psychoeducation; and technology-based counseling as part of a comprehensive program with nontechnology-based components.

The design of the included studies varied (refer to Table S2 in Multimedia Appendix 1). Studies applying a quantitative descriptive design mostly focused on users’ demographics, topics discussed and advice provided in counseling sessions, and satisfaction with services. Few of the included studies exclusively focused on implementation, and we found process evaluation reports [79,80] related to 2 interventions. Furthermore, the publication type of the included reports varied greatly, as we aimed to depict the broad spectrum of interventions. In addition to research reports, we identified abstracts, letters to the editor, and practical project reports.

The interventions are displayed in Table S2 in Multimedia Appendix 1 and characteristics of the included studies are described in Table S3 in Multimedia Appendix 1.

Results of Individual Sources of Evidence

Table 2 provides an overview of the data extracted from the included studies.

The label ("✓") indicates the presence of data without any information on content or scope. None of the interventions included reported data in all categories. The information available ranged from 1 to 7 conceptual categories for each intervention (also refer to the analysis matrices in Multimedia Appendix 1).
## Table 2. Overview of reported data on conceptual categories.

<table>
<thead>
<tr>
<th>Intervention</th>
<th>References</th>
<th>Acceptability</th>
<th>Adoption</th>
<th>Appropriateness</th>
<th>Feasibility</th>
<th>Fidelity</th>
<th>Implementation cost</th>
<th>Penetration</th>
<th>Sustainability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admiral Nurse Helpline</td>
<td>[29-32]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ADS Helpline</td>
<td>[33]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alzheimer Helpline</td>
<td>[34,35]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alz i-Connect</td>
<td>[36]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CANDIDbic</td>
<td>[37]</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Care Consultation</td>
<td>[38]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Care Consultation Plus</td>
<td>[38]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Care Consultation/Care Consultation Plus^c</td>
<td>[38]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coyne^d comparator</td>
<td>[39]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coyne^d experimental</td>
<td>[39]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coyne^c,d</td>
<td>[39]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helpline Alz Ass East Massa^c</td>
<td>[40]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Natale^d</td>
<td>[41]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ODCC^f</td>
<td>[42]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sabat^d</td>
<td>[43]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salfi^d nonanonym</td>
<td>[44-46]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salfi^d anonym</td>
<td>[44-46]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salfi^c,d</td>
<td>[44-46]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FITT-C^g</td>
<td>[47-53]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FITT-D^h</td>
<td>[54]</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NVAMP^d</td>
<td>[55]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICSS^j</td>
<td>[56-61]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>InformCare</td>
<td>[62-64]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Link2Care</td>
<td>[65]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Online Coaching</td>
<td>[66]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Cola^d</td>
<td>[67]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laver^d</td>
<td>[68,69]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RCTM^k</td>
<td>[70-73]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dementelcoach</td>
<td>[74-79]</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The mode of increased service demands or restrictions in the living conditions referring to aspects that need to be addressed, for example, connecting individuals to helpful information, providing access commitment to helping families affected by dementia, externally decision. The motives of organizations included their commitment to helping families affected by dementia, connecting individuals to helpful information, providing access to support services, and providing support. Some authors described organizational motives in more general terms by referring to aspects that need to be addressed, for example, increased service demands or restrictions in the living conditions of people with dementia in the community. The mode of decision was characterized in different ways: as a response to developments in the setting, such as mobility restrictions or the increasing use of the internet in the target groups; as consent to participate at the organizational level; or as permission sought and obtained to implement the program. External reasons for the uptake of the specific intervention were evidence of the effectiveness of the intervention found in the literature or evidence indicating that the previously used mode of delivery needed to be adapted. Internally motivated uptake is based on the development, modification, or advancement of one’s own interventions.

### Acceptability

We defined acceptability as the perception among implementation stakeholders of technology-based counseling that the intervention is agreeable, palatable, or satisfactory [17]. Within the data related to the conceptual category (reported for 20 interventions [29-36,38,40,43-69,74-80]; refer to Table 2), we identified the dimensions 

### Appropriateness

Appropriateness is understood as the perceived fit [17] of technology-based counseling for stakeholders, the setting, and the problems addressed. We defined the dimensions overall compatibility with stakeholder needs, tailoring to individuals, skills and instruments for enhancing fit, and concepts for fit for this conceptual category and found extensive information for all interventions [29-80]. If assignment of the quotes to experimental and comparator interventions was not possible, we assumed that the information was applicable for both interventions (refer to Table 2).

The dimension overall compatibility with stakeholder needs comprises information on accessibility, availability, tailoring to consumer groups, and usefulness of service. In the area of accessibility, the ways to access were described, ranging from the use of a single technology to multiple ways via email, telephone, and videoconferencing, or in combination with home visits. The availability of counseling in terms of service hours, for instance, permanently or during regular working hours on weekdays, was discussed against the background of availability when needed or in times of crisis. From the perspective of organizations, providers, and consumers, technology-based counseling was viewed as a mode of delivery that can solve logistical issues such as making appointments or long-distance travel. People who are homebound can access support without leaving their homes, and caregivers do not have to arrange substitute care. The limitations of remote delivery, such as the loss of context or consumers’ different capacities for using technology, were discussed from the perspective of the provider.
Implementation Cost

On the basis of Proctor et al [17], we defined the conceptual category **implementation cost** as the cost impact of an implementation effort reported from the perspective of a provider or the providing institution (reported for 3 interventions [36,62-64,74-79]; refer to Table 2). We identified the dimensions cost impact of delivery because of complexity of intervention, cost impact of implementation because of complexity of implementation strategy, and cost impact because of varying complexity of settings. Data on the first dimension comprised the costs of delivering the complex intervention and its financing through previously paid travel costs. The costs of the implementation strategy were illustrated by the impact of existing and lacking financial resources for staffing on the implementation process. Within the third dimension, failed expansion or implementation because of a lack of resources was exemplified.

Penetration

Within the conceptual category of **penetration**, defined as a step of integrating the technology-based counseling intervention into the service setting ([17] reported for 11 interventions [29-37,40,62-65,67,74-80]; refer to Table 2), we identified 3 dimensions: collaboration with stakeholders, access to the service, and spread. Data from interventions describe cooperation with stakeholders to implement the intervention, initiated either by the implementing organization itself or jointly through cooperation with other stakeholders in the setting. Access to the service occurred through referrals from other stakeholders in the setting, for example, physicians, or through information from other sources, such as telephone books or newspapers. From the perspective of the implementing organization, the degree of utilization of the intervention by different professional groups (eg, physicians, social workers, and nurses) was described. The level of spread was reported from the perspective of the implementing organization as well as the setting and is often reflected as the regional spread of interventions, for example, at the national or international level.

Sustainability

Following Proctor et al [17] and consequently Steckler et al [81], we understand **sustainability** (reported for 13 interventions [29-38,40,42,47-53,62-67]; refer to Table 2) as the final phase of the diffusion process in which innovations become entrenched in organizations. We were able to identify specifications of the data in 3 dimensions: routinization, passage, and incorporation. These dimensions were mainly reported from the perspective of the implementing organization; once, the perspective of the administration was also taken. The data on the dimension routinization provided information on the permanence and the degree to which the intervention was established, especially related to the number of versions of the intervention developed, the stakeholders involved, or the period from the start of the program. The duration varied greatly, ranging from a recent introduction to a multiyear build-up with many contacts. Statements were found in the interventions indicating maintained procedures, with the (planned) transition to expand the intervention often explained from the perspective of the organization with the aim of maintaining the intervention. An
administrative perspective was also taken, referring to the discontinuation of support after the end of the research project and, thus, the termination of the program. For example, this dimension was clarified by the integration of another target group or expansion to another region. Incorporation, for example, the final integration into (existing) organizational structures with the aim of maintaining the intervention, was mentioned, describing the dissemination of the service within existing structures.

**Figure 2.** Net diagram on reported data of conceptual categories.

**Graphical Presentation of Synthesized Data**

*Figure 2* shows the graphical synthesis of the data. Data on the conceptual categories of *implementation success* were sought for 27 interventions (blue line). The number of interventions for which data were extracted is indicated by the red line. While the categories *appropriateness* and *acceptability* are largely covered, substantial parts of the other areas remain unconsidered (refer to *Figure 2*).

**Discussion**

**Principal Findings**

We aimed to review the knowledge about the implementation success of technology-based counseling interventions for people with dementia or their caregivers. In our scoping review, we included 52 publications that reported 27 interventions. Interventions were heterogeneous and ranged from single counseling interventions, such as helplines, to counseling as part of multicomponent programs. To operationalize *implementation success*, we used the 8 outcomes of the conceptual framework for implementation outcomes [17] as conceptual categories. Only a few studies evaluated the implementation. Reporting on implementation outcomes was found to be fragmentary, and the comprehensiveness of the information varied widely. Overall, the focus of reporting was on the outcomes of *appropriateness* and *acceptability*, which may be because great efforts were made to adapt the interventions to a vulnerable target population.

As our data show, reporting on *acceptability* is inconsistent in terms of the perspective taken: data reports on measures to promote acceptability, but only from the perspective of the provider. At the same time, reporting on impact from this perspective was underrepresented. As already discussed in the literature, there are difficulties in unifying the wealth of perspectives in the context of implementation research [16].

The information reported on the dimension of *adoption* illustrates the importance of the fit between organizational motives and the type of intervention chosen for successful implementation. Little data are available on the mode of decision but differences in organizational culture can be identified that may influence the success of implementation efforts. We found data indicating that decisions from administrative stakeholders had an impact on sustainability by limiting the duration of the implementation of an intervention. Increasing the administrators’ enthusiasm for implementing the intervention by promoting familiarity with the specific intervention and using the effect of name-brand recognition may facilitate long-term commitment [82]. In some cases, the uptake of a specific intervention was based on evidence of its effectiveness, and there is a need for further effectiveness trials to expand the evidence base for decision-making regarding the implementation of technology-based counseling interventions.

The data provided on *appropriateness* comprised the largest amount of information extracted for categories and document the efforts undertaken to fit the target population. In addition to general measures to enhance the perceived fit of individualized support services, providers’ skills and instruments, as well as concepts applied by individual providers,
were described. Tailoring and personalizing counseling services to individuals’ needs has been associated with the usefulness of information and support [55-65]. The benefits and limitations of using technology for delivering counseling are discussed against the background of statements by consumers who would have preferred a different mode of communication with counselors [36]. On the basis of participants’ attrition, González-Fraile et al [9] reported that remote support or training interventions appear to be less acceptable to informal caregivers of people with dementia than control interventions, which may limit their applicability in community settings. Further research is needed to determine whether services that are accessible both face to face and technology-based can provide appropriate accessibility and improve the perceived fit of the target populations.

Information on feasibility comprised the successful implementation of interventions. Although factors impeding feasibility, such as legal issues and technical challenges, were reported, we found no information on failed or unsuccessful implementation. Barriers to the implementation of eHealth interventions described in the literature are, among others, the lack of digital literacy in the target population and staff’s uncertainties and insecurities about their coaching competences [83]. According to Proctor et al [17], the concept of feasibility is typically “invoked retrospectively as a potential explanation of an initiative’s success or failure.” Thus, a more comprehensive reporting of factors promoting or impeding feasibility may inform the implementation of interventions in future projects and may contribute to increasing the implementation success of technology-based counseling interventions.

Ways of formalizing the interventions to ensure fidelity in the delivery of interventions were mentioned, but manuals or guidelines were not made accessible along with publications. In addition, measurements to ensure fidelity were described for some interventions, but the results of assessments, as well as details on aspects where deviations occurred, were not reported. After critically reviewing the literature on the use of fidelity implementation frameworks in early intervention, Lemire et al [84] also stated gaps in defining and assessing implementation fidelity. Drawing on preexisting conceptualizations, the authors proposed a definition of fidelity that comprises the 4 components: adherence, exposure, quality, and participant responsiveness [84].

The cost impact of implementation efforts was rarely reported for the included interventions. Factors that influence the costs are the complexity of the specific intervention, the strategy used for implementation, and the delivery setting [17]. Despite the costs incurred in setting up the technical infrastructure, the costs for remote delivery were lower than when counseling was provided face to face [36,68,69]. Owing to the lower costs, eHealth interventions are considered suitable for widespread implementation [16]. The provision of information on implementation costs is essential to compare the cost impact of different interventions and to inform decisions regarding the uptake of a specific intervention [17].

The data reported on penetration often indicates access to the service in multiple ways, which seems to match the preferences and capabilities of consumers. As reported by Jelly et al [85], caregivers use dementia support services primarily when services are able to meet consumers’ individual needs. However, it is important to keep in mind that, from an organizational point of view, these extended access options need to be served simultaneously. In particular, cooperation with other stakeholders seems to be central to integrating counseling services into a service setting, but this was only highlighted by some authors in the included publications.

The sustainable anchoring of diffusion processes is described as a difficult phase in the implementation process of support services for caregivers of people with dementia. The reasons for this include a lack of understanding of the barriers to sustainable implementation in practice and a lack of long-term funding [86]. The problem is substantiated in that, as long as researchers focus on measuring the effectiveness of the interventions, the potential goal is not fully realized. However, there are models that can support this sustainable implementation [87].

There are several theoretical approaches, such as generalized theories, models, or frameworks, that address different aspects of implementation [88]. Frameworks “describe more loosely structured constellations of theoretical constructs... or prescriptive approaches for accomplishing implementation goals” [88]. By providing clarity in terms and definitions [88], frameworks contribute to shared language in implementation research. There are different types of frameworks focusing on processes or determinants or evaluations [88]. Evaluation frameworks, such as the conceptual framework for implementation outcomes introduced by Proctor et al [17] and the Reach, Efficacy, Adoption, Implementation, Maintenance (RE-AIM) planning and evaluation framework [89,90], offer guidance on identifying results that can be used to evaluate implementation efforts [88]. While the RE-AIM framework describes outcomes across 5 domains (reach, effectiveness, adoption, implementation, and maintenance) [90], Proctor et al [17] present the concept of 8 implementation outcomes, which are differentiated from service system and treatment outcomes. Implementation outcomes are defined as “the effects of deliberate and purposive actions to implement new treatments, practices, and services” and are reported from different levels of analysis (eg, the provider or the consumer perspective) [17]. Serving as conceptual categories of the implementation success [17], these outcomes provided the appropriate approach to operationalize the object of interest—the implementation success of technology-based counseling interventions in dementia—in our review.

To increase the clarity of terminology used in implementation research, Proctor et al [17] proposed the definitions of 8 conceptually distinct implementation outcomes as a “working taxonomy,” including different aspects of implementation success and thus creating a comprehensible framework. The use of outcomes as conceptual categories was sometimes challenging in our case. The mapping of the extracted data, in particular, was sometimes difficult because of the conceptual
similarity of some outcomes, for example, penetration and sustainability, and the inconsistent use of terminology found in
the literature. When determining the levels of analysis, we sometimes included additional perspectives, as described by
Proctor et al [17].

Altogether, the lack of process evaluation studies, the fragmented reporting, and the unclear use of terms and concepts
made it impossible to determine the extent of implementation success of technology-based counseling interventions in
dementia care. Because of the impaired comparability of data, we were not able to assess how the different types of
interventions affect the conceptual categories of acceptability, adoption, appropriateness, feasibility, fidelity, implementation
cost, penetration, and sustainability. For instance, we found data on appropriateness for all interventions, but the
consumer-perceived usefulness of services was referred to for only 5 interventions. These 5 interventions comprised helplines
that provided counseling via telephone [29-32,55] and web-based psychosocial interventions that provided information,
communication, and counseling [56-65]. The reported information does not allow any conclusions to be drawn on how
the types or components of interventions have an impact on implementation success. Gaining further insight into this issue
is important for developing future interventions that can be implemented successfully.

Adherence to the framework for developing and evaluating complex interventions [6] may help overcome these problems,
as proper process evaluation and exploration of conditions for implementation are recommended. The update to the Medical
Research Council guidance states that “[e]arly consideration of implementation increased the potential of developing an
intervention that can be widely adopted and maintained in real-world settings” [6] and thus can increase the success of
implementation efforts. In addition, the conceptual clarity of the terms and concepts used in implementation research is
needed to enhance transparency. This can be achieved by applying theoretical approaches that are “encapsulated as
generalized theories, models, or frameworks” [88]. The consistent use of terms not only creates clarity but also forms
the basis for better reporting on the success of implementation efforts, as Lengnick-Hall et al [91] proposed as the first of 6
practical recommendations for improved implementation outcomes reporting.

Strengths and Limitations
We followed a theory-driven approach to review the available evidence on implementation success. As we examined a broad
topic with evidence emerging from studies in various designs, a scoping review proved appropriate. We performed a
comprehensive and methodologically rigorous systematic literature search and included a variety of technology-based
counseling interventions for people with dementia and their informal caregivers. Differentiating counseling from
interventions focusing on education and information or from psychotherapeutic approaches brought challenges that we
overcame through intensive discussions in the review team. Although we were able to include a considerable number of
publications, it was not possible to make reliable statements about the implementation success of technology-based
counseling interventions in dementia because of the inconsistent database as well as the heterogeneity in terminology and
concepts.

Conclusions
We applied 8 conceptually distinct categories to operationalize the implementation success of technology-based counseling
interventions for people with dementia and their informal caregivers. We found considerable data for the categories
appropriateness and acceptability, and limited data on sustainability, feasibility, adoption, penetration, fidelity, and
implementation cost. There is an imbalance in the scope and depth of the reported data on the conceptual categories, and the
data extracted from the included publications only partially covered the concept of implementation success.

This highlights the need for a systematic evaluation accompanying the implementation of technology-based
counseling interventions in the context of dementia. Adherence to guidelines for the development and evaluation of interventions
and to guidelines or recommendations for reporting conceptualizations, measurements, and results on implementation outcomes is needed to expand knowledge on the effectiveness of implementation efforts and may foster the
implementation of complex interventions in diverse contexts.

Acknowledgments
This study was funded by the German Federal Ministry of Education and Research (grant number 01KG2110). The funder had
no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation,
review, or approval of the manuscript; or decision to submit the manuscript for publication. We acknowledge the financial support
of the Open Access Publication Fund of the Martin Luther University Halle-Wittenberg.

The authors thank Fabian Wilde, MSc, Leipzig (Germany), for supporting them in the selection process of the studies.

Authors’ Contributions
DB, JW, JH, GM, and AB conceived the idea of the manuscript and made substantial contributions to this scoping review. JH
and AB conducted the literature search. DB, JW, JH, and AB conducted study selection and data extraction. DB, JW, JH, and
AB conceptualized and operationalized the object of interest. DB and JW conducted the data analysis and synthesis. DB and JW
were involved in the drafting of the manuscript. DB, JW, JH, GM, and AB revised the manuscript for important intellectual
content. All authors read and approved the final manuscript.

https://aging.jmir.org/2024/1/e51544
Conflicts of Interest
None declared.

Multimedia Appendix 1
Description of included intervention programs, description of studies reporting on included intervention programs, and analysis matrices.

[DOCX File, 181 KB - aging_v7i1e51544_app1.docx]

Multimedia Appendix 2
Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist.

[DOCX File, 89 KB - aging_v7i1e51544_app2.docx]

References


44. Salfi J. Seeking to understand telephone support for dementia caregivers: a qualitative case study. McMaster University. 2004 Sep. URL: https://macsphere.mcmaster.ca/bitstream/11375/7569/1/fulltext.pdf [accessed 2022-12-05]


46. Spilsbury K. Commentary: Telephone support met the perceived needs of dementia caregivers for convenient access to information, referral, and emotional support. Evid Based Nurs 2006 Jul;9(3):94. [doi: 10.1116/ebn.9.3.94] [Medline: 16865845]


Abbreviations

PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

RE-AIM: Reach, Efficacy, Adoption, Implementation, Maintenance

©Dorothee Bauernschmidt, Janina Wittmann, Julian Hirt, Gabriele Meyer, Anja Bieber. Originally published in JMIR Aging (https://aging.jmir.org), 25.01.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Review

Patient and Public Involvement in Technology-Related Dementia Research: Scoping Review

Pippa Kirby1,2,3, BA, MSc; Helen Lai2,3, BSc, MRes; Sophie Horrocks2,4, MA, MSc; Matthew Harrison2,4, MA, MEng; Danielle Wilson2,3, BSc; Sarah Daniels2,3, MSc; Rafael A Calvo5, PhD; David J Sharp2,3, BA, MBBS, PhD; Caroline M Alexander1,6, MCSP, MSc, PhD

1Department of Therapies, Imperial College Healthcare NHS Trust, London, United Kingdom
2UK Dementia Research Institute, Care Research and Technology Centre (UK DRI CR&T), London, United Kingdom
3Department of Brain Sciences, Faculty of Medicine, Imperial College London, London, United Kingdom
4Helix Centre, Institute of Global Health Innovation, Imperial College London, London, United Kingdom
5Dyson School of Design Engineering, Imperial College London, London, United Kingdom
6Department of Surgery and Cancer, Imperial College London, London, United Kingdom

Corresponding Author:
Helen Lai, BSc, MRes
UK Dementia Research Institute
Care Research and Technology Centre (UK DRI CR&T)
9th Floor, Sir Michael Uren Hub, White City Campus, Imperial College London
86 Wood Lane
London, W12 0BZ
United Kingdom
Phone: 44 20 7594 9755
Email: h.lai18@imperial.ac.uk

Abstract

Background: Technology-related research on people with dementia and their carers often aims to enable people to remain living at home for longer and prevent unnecessary hospital admissions. To develop person-centered, effective, and ethical research, patient and public involvement (PPI) is necessary, although it may be perceived as more difficult with this cohort. With recent and rapid expansions in health and care–related technology, this review explored how and with what impact collaborations between researchers and stakeholders such as people with dementia and their carers have taken place.

Objective: This review aims to describe approaches to PPI used to date in technology-related dementia research, along with the barriers and facilitators and impact of PPI in this area.

Methods: A scoping review of literature related to dementia, technology, and PPI was conducted using MEDLINE, PsycINFO, Embase, and CINAHL. Papers were screened for inclusion by 2 authors. Data were then extracted using a predesigned data extraction table by the same 2 authors. A third author supported the resolution of any conflicts at each stage. Barriers to and facilitators of undertaking PPI were then examined and themed.

Results: The search yielded 1694 papers, with 31 (1.83%) being analyzed after screening. Most (21/31, 68%) did not make clear distinctions between activities undertaken as PPI and those undertaken by research participants, and as such, their involvement did not fit easily into the National Institute for Health and Care Research definition of PPI. Most of this mixed involvement focused on reviewing or evaluating technology prototypes. A range of approaches were described, most typically using focus groups or co-design workshops. In total, 29% (9/31) described involvement at multiple stages throughout the research cycle, sometimes with evidence of sharing decision-making power. Some (23/31, 74%) commented on barriers to or facilitators of effective PPI. The challenges identified often regarded issues of working with people with significant cognitive impairments and pressures on time and resources. Where reported, the impact of PPI was largely reported as positive, including the experiences for patient and public partners, the impact on research quality, and the learning experience it provided for researchers. Only 4 (13%) papers used formal methods for evaluating impact.

Conclusions: Researchers often involve people with dementia and other stakeholders in technology research. At present, involvement is often limited in scope despite aspirations for high levels of involvement and partnership working. Involving people with dementia, their carers, and other stakeholders can have a positive impact on research, patient and public partners, and...
Wider reporting of methods and facilitative strategies along with more formalized methods for recording and reporting on meaningful impact would be helpful so that all those involved—researchers, patients, and other stakeholders—can learn how we can best conduct research together.

**KEYWORDS**

dementia; technology; patient and public involvement and engagement; co-design; coproduction

**Introduction**

**Background**

Worldwide incidence of dementia is increasing. In the United Kingdom alone, there are approximately 1 million people living with dementia, with this figure expected to double by 2050 [1]. The total cost of care for people with dementia in the United Kingdom in 2019 was £34.7 billion (US $44.1 billion), with an expected increase to approximately £94 billion (US $119.5 billion) by 2040 [2]. Technology is increasingly cited as a means of supporting people with dementia and their formal and informal carers and reducing some of this economic burden. “Digitally enabled care” is a core component of the National Health Service Long Term Plan [3]—it is felt that technology has the potential to facilitate aging in place and reduce unplanned hospital admissions, with consequent economic benefits as well as improved health outcomes and quality of life [4,5]. Smart home systems, assistive technology devices, and other technologies are being developed with aims including supporting safety in people’s homes; enabling early detection of deterioration or ill health; supporting activities of daily living; and facilitating access to treatment, leisure activities, or social participation [5-8].

Despite rapid advancements in technology, the implementation of health and care–related technology for people with dementia has been slow, and there is increasing recognition of the many challenges in this area [9-11]. These challenges include ethical issues regarding privacy, autonomy, safety, and trust and the risk of creating or exacerbating health-related bias and inequality [12-16]. Researchers and technology developers must also understand the complex and changing needs of individual circumstances—there is a need for research to center on the person and their support network rather than on the technology itself if it is to be successful [17]. Therefore, understanding users’ perspectives is fundamental if we are to develop technologies that are acceptable, effective, and ethical [5,10,18,19]. One way to achieve this is through patient and public involvement (PPI).

PPI describes a partnership between patients, the public, and researchers in the research process itself. It is often described as research conducted “with” or “by” service users rather than research “about” or “for” them [20]. In addition to being seen as an ethical imperative, PPI aims to improve the efficiency and value of health research, recognizing that those with lived experience of health conditions or services will bring knowledge and experience that may increase the relevance of studies, improve recruitment and retention of participants, and improve dissemination of research findings [20-22]. PPI is now seen as an essential part of health and social care research—the Health Research Authority strongly advises PPI because of its likelihood of improving research quality and addressing the Research Ethics Committee’s key considerations [23]. Stakeholder engagement is a key part of the guidance from the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use [24], and the National Institute for Health and Care Research (NIHR) makes it a condition of research funding [21]. The NIHR describes different approaches to involvement with increasing levels of power and influence for members of the public, from consultation (least power) to coproduction and user controlled (most power) [20].

PPI in dementia-related research has been gathering pace in recent years. Historically focusing more on the involvement of carers or other stakeholders, this has changed with more studies involving people with dementia themselves [25,26]. It is now well established that this supports and promotes a person-centered model of health care [27-31]. PPI should be conducted in a manner that promotes equality, diversity, and inclusion [20]. The NIHR emphasizes the need for researchers to enable the involvement of underrepresented groups and adapt their PPI approaches and activities to ensure accessibility for all groups affected by the project [32]. When planning and carrying out PPI in dementia research, this means the consideration of all groups affected by aging and dementia from diverse ethnic, racial, linguistic, geographic, and socioeconomic backgrounds. The principle of stakeholder involvement is not unique to PPI. To understand the principles of terms such as “co-design” and “coproduction” within PPI, it is important to appreciate the context in which these terms have developed beyond just the health care sector. Within technology innovation, there has been a steady and increasing emphasis over the past 50 years on ensuring that a “human-centered” approach is taken to developing a new product or service [33]. Human-centered design emphasizes the need for fostering deep empathy with the people one is designing with, bringing end users into the design process as early as possible. Co-design can be a method of human-centered design. Co-design also stems from the 1970s, from a Scandinavian movement of participatory design, in which scientists, technologists, and design researchers acknowledged that “the people destined to use the system [must] play a critical role in designing it” [34]. Wider adoption of these human-centered design approaches has been seen in the last 15 years with methodologies such as the Design Council’s “Double Diamond” [35] helping visualize this iterative approach to innovation and wider adoption across nondesigners. Considering the context of technology within health and social care, it is not surprising that practitioners from health and social care, it is not surprising that practitioners from health and social
care, design, and technology research have found themselves discussing what best practice should look like and what approaches or methods might facilitate meaningful innovation [36]. Regardless of the background, researchers across these disciplines agree on the need to move from a patient-centered or user-centered approach to a “co-production” approach in which users not only are observed or consulted but also work jointly as partners, with mutual respect and understanding of each other’s different knowledge and experiences and the contributions they can make [21,37,38]. The NIHR outlines 5 key principles of coproduction as part of a research project (Textbox 1).

Textbox 1. National Institute for Health and Care Research “Guidance on co-producing a research project”—key principles.

1. Sharing of power—the research is jointly owned and people work together to achieve a joint understanding
2. Including all perspectives and skills—making sure the research team includes all those who can make a contribution
3. Respecting and valuing the knowledge of all those working together on the research—everyone is of equal importance
4. Reciprocity—everyone benefits from working together
5. Building and maintaining relationships—an emphasis on relationships is key to sharing power” [32]

Despite the recognition of the value of PPI and the recommendation of coproduction approaches [10,19], patient or other stakeholder involvement in technology-related dementia research is known to be variable in breadth and depth and sometimes absent altogether [11]. Older reviews show that the involvement of people with dementia has usually been as passive participants to be observed or at most as a group to consult but without any sharing of decision-making power [19,39,40]. A review of the literature published between 2011 and 2017 by Suijkerbuijk et al [41] demonstrated that, although there has been an increase in the involvement of people with dementia in technology research, reporting on the methods, barriers, facilitators, and impact remains minimal, making progress in this field challenging. This mirrors issues with PPI reporting (especially of impact) in the wider field of dementia research [25,42,43]. Given the increased attention that PPI has received in recent years as well as the rapid advances in technology-related health research, we anticipated that there would be many more papers published in the period from 2017 to 2022 worthy of review. In addition, the review by Suijkerbuijk et al [41] included papers with a broad range of methodologies to cover the concept of “involvement,” including the involvement of people with dementia as participants in qualitative research. To our knowledge, no review to date has explored the specific concept of PPI in technology-related dementia research.

Objectives

Therefore, the objectives of this scoping review were as follows:

1. To describe the approaches to PPI used to date in technology-related dementia research, exploring who is involved, when, and how,
2. To describe the reported barriers to and facilitators of effective PPI in this area, and
3. To examine and report on the impact of PPI in this area.

Methods

Review Type

To gather the available literature in this area, a scoping review was conducted. Scoping reviews are often used in preference to systematic reviews in cases in which the body of literature is likely to be large and heterogeneous and to answer broad questions (such as “what is known about this concept?”) [44]. They are a useful way to map out the evidence, as opposed to systematic reviews, which often bring together literature on a particular subject with a more defined question, for example, about the efficacy of interventions [45]. The PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) guidelines [44] were followed to ensure appropriate reporting.

Search Strategy and Eligibility Criteria

A search strategy was developed and used a search string consisting of words related to dementia; technology designed to support health, care, or well-being; and PPI. Knowing that the terminology used varies considerably, definitions were kept broad, in particular of “patient and public involvement,” adapting and building on existing search strings from previous reviews [11,19,25,41,42]. Our definition of technology was similarly broad. Assistive technology may be described as “products or systems that support and assist individuals with disabilities, restrict mobility or other impairments to perform functions that might otherwise be difficult or impossible” [46]. We included any type of assistive technology as well as, more broadly, any technology that could be deemed to be a part of technology-enabled care (such as telehealth systems, telecare, telemedicine, and self-care apps) [47]. Inclusion criteria were developed (Textbox 2). PPI activities do not usually require ethics approval [20], yet we did not exclude those who sought ethics approval so as to ensure that we captured a range of approaches.
Textbox 2. Inclusion and exclusion criteria.

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Research about dementia (any type) or mild cognitive impairment</td>
</tr>
<tr>
<td>• Research focused on technology designed to support the health, care, or well-being of people with dementia or their carers</td>
</tr>
<tr>
<td>• Research describing ways in which patients or other stakeholders were actively involved in the research process itself (not only as research participants)</td>
</tr>
<tr>
<td>• Full text available in English</td>
</tr>
<tr>
<td>• Any publication date up to the end of 2022</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Dementia only mentioned incidentally (eg, primary focus was Parkinson disease)</td>
</tr>
<tr>
<td>• Technology in which target beneficiaries are not people with dementia, family or carers (eg, web-based education programs for health care workers)</td>
</tr>
<tr>
<td>• Studies in which the patients or stakeholders are positioned as research participants only (eg, participants in a qualitative study) and are not actively involved in conducting the research</td>
</tr>
<tr>
<td>• Reviews</td>
</tr>
<tr>
<td>• Opinion pieces</td>
</tr>
<tr>
<td>• Conference abstracts</td>
</tr>
</tbody>
</table>

Data Sources and Charting Process

The search was conducted in 4 databases: MEDLINE, PsycINFO, Embase (using Ovid), and CINAHL (using EBSCO). All papers published until the end of 2022 were included. Abstracts had to be available in English, and opinion pieces and reviews were excluded (refer to Multimedia Appendix 1 for the full MEDLINE search string). The search was last conducted in January 2023. References were exported to EndNote (Clarivate Analytics) and then to Covidence (Veritas Health Information) for screening. After the removal of duplicates, 2 reviewers (PK and HL) screened the titles and abstracts against the eligibility criteria. The full texts were then further screened for eligibility. The 2 reviewers then independently charted the data from the included studies using a predesigned extraction table. For the first 10 papers, detailed discussions were held to clarify interpretations of PPI. Subsequent discussions were held to reach a consensus where required. A third author (CMA) was available if a consensus was not reached. As the purpose of this review was to provide an overview of existing evidence regardless of quality, no formal appraisal of methodological quality was conducted, in line with guidance [45]. Facilitators of and barriers to effective PPI were grouped and analyzed by the first author to draw out themes, which were then refined in discussion with the other authors. The impact of PPI, where described, was summarized and categorized into impact on the study, impact on the patient and public partners, and impact on the research team.

Initial database searching identified 1689 records, with an additional 5 found through hand searches following references from papers identified in the initial search. After removal of 695 (41%) duplicates, the remaining 999 abstracts were screened. Most of these (915/999, 91.6%) did not meet the eligibility criteria (were not about dementia, involvement in research, or technology). Determining whether papers described active involvement in the research process or merely involvement as participants was frequently unclear from the abstracts alone, and the authors erred on the side of inclusion here, in line with guidance. When analyzing full texts (84/999, 8.4%), not meeting the “involvement in research” criteria was the most common reason for exclusion (25/53, 47% of the papers excluded at this stage). A total of 31 papers were included in the scoping review. Figure 1 shows the flow of information for this process.
Results

Study Characteristics

Of the 31 papers included in the review (Table 1), most (n=18, 58%) were published between 2020 and 2022. Only 13% (4/31) were published before 2010. Most papers were authored by groups from multiple disciplines, for example, authors from design engineering backgrounds and health sciences and health care professionals. A total of 13% (4/31) of the studies included patient or public partners as coauthors [17,49-51]. In total, 21 of the studies originated in Europe, including 10 (48%) from the United Kingdom. Others were from Canada (7/31, 23%) and Australia (3/31, 10%), and 1 of the studies (1/31, 3%) included discussions of collaborations with groups in Ireland, Hong Kong, Brazil, and India [52].

Table 1 also outlines the stakeholders involved, the stage or stages of the research they were involved in, and the methods used for this involvement. When determining which stage of research stakeholders were involved in, the authors mapped involvement onto NIHR stages of research in which PPI might typically take place (eg, “design of the research”). Where Table 1 states “did not fit model,” this was because participants were positioned as both coresearchers and research participants. This is explored further in the following sections (Table 2).

A range of technologies were described with varied purposes (Textbox 3), apart from in Liddle et al [17], which did not focus on any one type.
Table 1. Study characteristics describing the technology, stakeholder involvement, stage of the research process in which the involvement took place, role of the stakeholder, and methods used to involve patients and the public.

<table>
<thead>
<tr>
<th>Study, year</th>
<th>Origin</th>
<th>Technology type and purpose</th>
<th>Stakeholders involved</th>
<th>NIHR stage of research in which PPI took place</th>
<th>Role of stakeholders involved</th>
<th>PPI methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Davies et al [53], 2019</td>
<td>United Kingdom</td>
<td>Website to provide information and peer and professional support for caregivers of people with dementia toward end of life</td>
<td>Research development group including 6 HCPs(^c), 2 members of a dementia charity organization, and 1 carer</td>
<td>Design of the research, undertaking and management of the research, and analysis of data</td>
<td>Refining the aims of the wider project and steering the project throughout</td>
<td>Steering group, group meetings, and 1:1 meetings</td>
</tr>
<tr>
<td>Kort and van Hoof [54], 2014</td>
<td>The Netherlands</td>
<td>Website with information about home modifications for people with dementia and their family caregivers</td>
<td>3 dyads (people with dementia+carer) and, separately, a group of 20 (carers or residential home staff)</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Observations, consultation rounds, and questionnaire</td>
</tr>
<tr>
<td>Schikhof et al [55], 2010</td>
<td>The Netherlands</td>
<td>Monitoring system in residential home for people with dementia to detect anomalies (eg, panicking and falls)</td>
<td>8 nursing home staff members and 13 family representatives (as proxies for people with dementia)</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Interviews, workshops for prototype testing, informal group meetings, and focus groups</td>
</tr>
<tr>
<td>Muñoz et al [56], 2022</td>
<td>Canada</td>
<td>Virtual reality game to support engagement in exercise for people with dementia</td>
<td>7 people with dementia or MCI(^d), 5 older adults without dementia, industry partners, and HCPs</td>
<td>Design of the research and undertaking and management of the research; elements did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Prototype testing and interviews (people with dementia or MCI and older adults), research group meetings (HCPs), and focus group (older adults without dementia, members of the research team, and industry representatives)</td>
</tr>
<tr>
<td>Eisapour et al [57], 2020, also with reference to Eisapour et al [58], 2018, and Eisapour [59], 2018</td>
<td>Canada</td>
<td>Virtual reality game to support engagement in exercise for people with dementia</td>
<td>HCPs and 3 people with dementia</td>
<td>One PPI representative involved in the main research team, presumed to be throughout; other elements did not fit model</td>
<td>Contributing to the iterative design process and involved in the main research team</td>
<td>Focus groups, observations, informal discussions in care home, and prototype testing; one member of the original focus group subsequently joined the research team</td>
</tr>
<tr>
<td>Hanson et al [60], 2007</td>
<td>Sweden</td>
<td>Home computer-based education and support program providing information, support tools, and exercises for people with dementia and their carers</td>
<td>7 people with dementia</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Group meetings to develop support program content and computer skills tuition for group members</td>
</tr>
<tr>
<td>Orpwood et al [61], 2004</td>
<td>Finland and others part of the ENABLE project</td>
<td>Various assistive technology devices: picture gramophone, calendar, tap monitor, lost object locator, gas cooker monitor, and night-light</td>
<td>Family carers (as proxies for people with dementia), paid carers, and older adults without dementia</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Informal group meetings and focus groups</td>
</tr>
<tr>
<td>Study, year</td>
<td>Origin</td>
<td>Technology type and purpose</td>
<td>Stakeholders involved</td>
<td>NIHR(^a) stage of research in which PPI(^b) took place</td>
<td>Role of stakeholders involved</td>
<td>PPI methods</td>
</tr>
<tr>
<td>------------</td>
<td>-------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>-------------------------------------------------</td>
<td>-------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Kort et al [62], 2019</td>
<td>The Netherlands</td>
<td>Various projects described: a smart pill box and real-time location systems (also a website as documented separately [51])</td>
<td>People with dementia and carers (past or current)</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Varied between projects: observations, consultations, storytelling, focus groups, and “thinking aloud” sessions</td>
</tr>
<tr>
<td>Hendriks et al [63], 2017, also with reference to Hendriks et al [64], 2014</td>
<td>Belgium</td>
<td>System for monitoring the mealtimes of people with dementia in a care home using sensors and data input by carers into the app</td>
<td>Industry representatives, academics and HCPs, professional carers, people with dementia, and informal carers</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Meetings, feedback sessions with HCPs or professional carers, integration of research team into daily life at care home, and group “mapping” sessions with people with dementia and carers</td>
</tr>
<tr>
<td>Orpwood et al [65], 2007</td>
<td>United Kingdom</td>
<td>Various technology projects: music player, video streaming of outside world scenes into the homes of people with dementia, conversation prompter for people with dementia, and “sequence support” tool for prompting ADLs(^c)</td>
<td>Varied across projects: people with dementia; carers; and other “user representatives,” including academics from the social sciences, engineering, and dementia organizations</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Varied across projects: user surveys, co-design workshops, observations, 1:1 user testing, and interviews</td>
</tr>
<tr>
<td>Savitch et al [66], 2012</td>
<td>United Kingdom</td>
<td>Website providing information about assistive technology for people with dementia</td>
<td>People with dementia and, separately, a steering group that also included 1 person with dementia</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process (people with dementia); steering group also provided input throughout</td>
<td>Interviews, focus groups, co-design workshops, and involvement in steering group meetings</td>
</tr>
<tr>
<td>Perkins et al [52], 2022</td>
<td>United Kingdom, Ireland, Hong Kong, India, and Brazil</td>
<td>Web-based delivery of CST(^d)</td>
<td>4 people with dementia, 4 family carers, 4 service managers, and 8 CST group facilitators from the United Kingdom and Hong Kong; additional stakeholders in India</td>
<td>Design of the research and undertaking and management of the research</td>
<td>Designing and developing a protocol (people with dementia, carers, service managers, and CST facilitators) and field-testing of the protocol and subsequently giving feedback following testing (CST facilitators)</td>
<td>Focus groups (web-based); CST facilitators then tested the protocol with people with dementia in 5 countries; interviews with CST facilitators following field-testing</td>
</tr>
<tr>
<td>Hwang et al [67], 2015</td>
<td>Canada</td>
<td>Animated videos for prompting people with dementia with ADLs</td>
<td>6 family carers</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Co-design workshops first to develop the concept and then refine the designs; 2 home visits for paper prototype evaluation</td>
</tr>
<tr>
<td>Oksnebjerg et al [68], 2019</td>
<td>Denmark</td>
<td>An app to support self-management for people with dementia, including a calendar and diary</td>
<td>4 people with dementia, 4 family carers, and 4 paid carers</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Co-design workshops</td>
</tr>
<tr>
<td>Study, year</td>
<td>Origin</td>
<td>Technology type and purpose</td>
<td>Stakeholders involved</td>
<td>NIHR(^a) stage of research in which PPI(^b) took place</td>
<td>Role of stakeholders involved</td>
<td>PPI methods</td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>-----------------------------</td>
<td>-----------------------</td>
<td>-------------------------------------------------</td>
<td>--------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>Hung et al [50], 2021</td>
<td>Canada</td>
<td>PARO, a commercially available robot seal that uses AI(^c) to support the social and emotional needs of the people with dementia interacting with it</td>
<td>5 “patient and family partners”</td>
<td>Undertaking and managing the research, analysis of data, and dissemination of research findings</td>
<td>Supporting data collection, thematic analysis of data, and authoring of the paper</td>
<td>Supporting data collection with some patients or particular settings, analysis (4 × 1-h thematic analysis group sessions), and coauthoring the paper</td>
</tr>
<tr>
<td>Rathnayake et al [37], 2021</td>
<td>Australia</td>
<td>A mobile health app that provides information about dementia, care strategies, and tips for managing ADLs</td>
<td>Carers, HCPs, and IT experts</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Web-based survey, interviews, and co-design workshops</td>
</tr>
<tr>
<td>Kowe et al [69], 2021, also with reference to Kowe et al [70], 2022</td>
<td>Germany</td>
<td>Sensor-based activity management system</td>
<td>6 family carers</td>
<td>Analysis of data</td>
<td>Supporting thematic analysis of interview data</td>
<td>30-min group analysis session or workshop</td>
</tr>
<tr>
<td>Daly Lynn et al [51], 2021, also with reference to Daly Lynn et al [71], 2019, and Daly Lynn et al [72], 2022</td>
<td>United Kingdom</td>
<td>Smart home system for people with dementia in supported living</td>
<td>Peer researchers: 7 older adults without dementia; steering group: including 2 people with dementia, 1 family carer, and 2 dementia organization employees</td>
<td>Undertaking and management of the research, analysis of data, and dissemination; steering group: detail not given</td>
<td>Peer researchers: conducting interviews with research participants and supporting data analysis; steering group: ensuring that the voice of older people was central to the project and coauthoring the paper</td>
<td>2-d training for peer researchers, conducting interviews jointly with a member of the research team, group thematic analysis session, and evaluation forms</td>
</tr>
<tr>
<td>Banbury et al [73], 2021</td>
<td>Australia</td>
<td>Virtual support program to provide information and peer support for carers of people with dementia</td>
<td>6 family carers</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Co-design workshops and group meetings (all virtual) following the Double Diamond approach</td>
</tr>
<tr>
<td>Fox et al [74], 2022</td>
<td>United Kingdom</td>
<td>A mobile health app that delivers memory tests throughout the day to monitor cognition changes</td>
<td>7 people with dementia, 7 family carers, and a PPI officer</td>
<td>Did not fit model</td>
<td>PPI officer as “proxy patient representative” in the research team; people with dementia and carers contributed to the iterative design process</td>
<td>Co-design workshops; PPI officer part of the main research group</td>
</tr>
<tr>
<td>Rai et al [75], 2020</td>
<td>United Kingdom</td>
<td>Virtual delivery of CST</td>
<td>People with dementia and family carers</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Consultation meetings and co-design workshops</td>
</tr>
<tr>
<td>Owens et al [76], 2020</td>
<td>Consortium spanning organizations in Europe and the United States</td>
<td>Remote monitoring and measurement technologies for people with dementia</td>
<td>People with dementia and family carers</td>
<td>Design of the research</td>
<td>Patient advisory board—supporting research planning and design and feedback on relevance and priorities</td>
<td>Provision of feedback on researchers’ literature review, group discussions, workshops, and other meetings of the patient advisory board</td>
</tr>
<tr>
<td>Study, year</td>
<td>Origin</td>
<td>Technology type and purpose</td>
<td>Stakeholders involved</td>
<td>NIHR(^a) stage of research in which PPI(^b) took place</td>
<td>Role of stakeholders involved</td>
<td>PPI methods</td>
</tr>
<tr>
<td>-------------</td>
<td>--------</td>
<td>-----------------------------</td>
<td>----------------------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>Stavropoulos et al [77], 2021</td>
<td>Greece, with involvement from multiple countries across Europe</td>
<td>A range of commercially available wearable devices</td>
<td>11 people with dementia and 10 carers from 11 countries across Europe</td>
<td>Design of the research</td>
<td>Reviewing devices and providing information to guide the design of future research, aiming to make it more relevant and accessible and improve participant experience</td>
<td>One-off 2.5-h session including presentations, roundtable discussions, hands-on experimentation, and voting</td>
</tr>
<tr>
<td>Liddle et al [17], 2022</td>
<td>Australia</td>
<td>No specific technology discussed—study explored factors related to engagement with technology for people with dementia and carers</td>
<td>15 people with dementia and carers (living experience expert reference group); 4 of them (2 people with dementia and 2 carers) were also members of the core research team and were listed as coauthors</td>
<td>Design of the research, undertaking and management of the research, analysis of data, and dissemination of research findings</td>
<td>Setting research priorities, supporting the design of interviews and developing the topic guide, thematic analysis of data, reflexivity sessions with the research team, and writing up of the study</td>
<td>Group sessions, discussions and meetings with reference group, group analysis sessions, and discussions of reflectivity</td>
</tr>
<tr>
<td>Hassan et al [78], 2017</td>
<td>United Kingdom</td>
<td>A range of commercially available wearable devices</td>
<td>&gt;30 in total: people with dementia or MCI, carers, people with early-onset dementia (aged &lt;65 y), and older adults without dementia</td>
<td>Design of the research</td>
<td>Contributing to research design (producing guidelines to optimize design and improve acceptability) and informing procurement decisions</td>
<td>Exploratory workshops in groups according to diagnosis (people with dementia+carers, people with early-onset dementia, and older adults without dementia), individual 1:1 meetings (people with MCI), and opportunities for stakeholders to try the devices at home</td>
</tr>
<tr>
<td>Jacklin et al [79], 2020</td>
<td>Canada</td>
<td>A wearable (wristband) for capturing movement-related behaviors (eg, falls, wandering, and agitation)</td>
<td>Indigenous community representatives, carers, community elders, and health and social care workers</td>
<td>Did not fit model</td>
<td>Informing the design of the research to ensure that culturally appropriate and inclusive methods are used</td>
<td>Community-based participatory research approach, focus groups (with carers), and preparation meetings with the Community Advisory Council to plan the research and ensure that appropriate methods are used</td>
</tr>
<tr>
<td>Ghafurian et al [80], 2022</td>
<td>Canada</td>
<td>App to support professional carers in communicating with people with dementia in nursing homes</td>
<td>17 professional carers and 1 nurse</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Survey, focus groups, and co-design workshops</td>
</tr>
<tr>
<td>Berge et al [81], 2022</td>
<td>Norway</td>
<td>Tablet-based music app primarily used as a relational tool to support positive interactions between people with dementia and carers</td>
<td>People with dementia or other psychiatric conditions, carers, and HCPs</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Observations, interviews, and 2 focus groups; user testing of a prototype with 4 older adults with dementia or other psychiatric conditions</td>
</tr>
<tr>
<td>Study, year</td>
<td>Origin</td>
<td>Technology type and purpose</td>
<td>Stakeholders involved</td>
<td>NIHRa stage of research in which PPIb took place</td>
<td>Role of stakeholders involved</td>
<td>PPI methods</td>
</tr>
<tr>
<td>------------</td>
<td>--------</td>
<td>----------------------------</td>
<td>-----------------------</td>
<td>-----------------------------------------------</td>
<td>-------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Shadarevian et al [49], 2020</td>
<td>Canada</td>
<td>Tablet for sharing family videos in hospital with people with dementia to support care provision, reduce social isolation, and reduce aggression</td>
<td>People with dementia, family carers, HCPs, and students</td>
<td>Undertaking and management of the research, analysis of data, and dissemination of research findings</td>
<td>Positioned as part of the research team throughout, guiding the research process; thematic analysis of interview data; creating a toolkit to support wider dissemination and adoption of the intervention; and coauthoring the paper</td>
<td>Monthly research project meetings, making videos, interviews, group thematic analysis session, manuscript reviewing and editing</td>
</tr>
<tr>
<td>Tiersen et al [82], 2021</td>
<td>United Kingdom</td>
<td>Smart home system for people with dementia in their own homes</td>
<td>People with dementia, carers, and HCPs</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Various across 9 substudies: semi-structured interviews, focus groups, co-design workshops, surveys, home visits, online group meetings, and observations</td>
</tr>
<tr>
<td>van der Roest et al [83], 2008</td>
<td>The Netherlands</td>
<td>Website with customized information for people with dementia and their carers about health care and welfare services</td>
<td>People with dementia, carers, and HCPs</td>
<td>Did not fit model</td>
<td>Contributing to the iterative design process</td>
<td>Workshop exploring user needs (people with dementia, carers, and HCPs), further co-design workshops (HCPs), prototype testing, and questionnaire (family carers)</td>
</tr>
</tbody>
</table>

aNIHR: National Institute for Health and Care Research.
bPPI: patient and public involvement.
cHCP: health care professional.
dMCI: mild cognitive impairment.
eADL: activity of daily living.
fCST: cognitive stimulation therapy.
gAI: artificial intelligence.

Table 2. Patient and public involvement (PPI) in the different stages of research as recommended by the National Institute for Health and Care Research [21] (n=31).
**Textbox 3. Types of technologies and their purposes.**

<table>
<thead>
<tr>
<th>Category</th>
<th>Purpose/Safety Alerts</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sensor monitoring systems (including smart home monitoring systems)</strong></td>
<td>• Safety alerts&lt;br&gt;• Monitoring (of health, activity, behavior, and cognition)</td>
</tr>
<tr>
<td><strong>Wearables</strong></td>
<td>• Safety alerts&lt;br&gt;• Monitoring (of health, activity, behavior, and cognition)</td>
</tr>
<tr>
<td><strong>Apps</strong></td>
<td>• Monitoring (of health, activity, behavior, and cognition)&lt;br&gt;Self-management&lt;br&gt;Exercises&lt;br&gt;Information provision or education&lt;br&gt;Supporting social interaction&lt;br&gt;Social and emotional well-being</td>
</tr>
<tr>
<td><strong>Miscellaneous assistive technology devices (not wearables and not app based, eg, gas cooker monitor or smart pill box)</strong></td>
<td>• Safety alerts&lt;br&gt;• Self-management&lt;br&gt;Leisure access&lt;br&gt;Supporting social interaction&lt;br&gt;Social and emotional well-being</td>
</tr>
<tr>
<td><strong>Websites</strong></td>
<td>• Information provision and education&lt;br&gt;Accessing peer support&lt;br&gt;Accessing professional support</td>
</tr>
<tr>
<td><strong>Videoconferencing platforms</strong></td>
<td>• Therapy delivery&lt;br&gt;Information provision and education&lt;br&gt;Accessing peer support</td>
</tr>
<tr>
<td><strong>Virtual reality games</strong></td>
<td>• Exercises</td>
</tr>
</tbody>
</table>

**Who Was Involved?**

Studies involving only 1 stakeholder group in their PPI activities were in the minority (5/31, 16%), and many (17/31, 55%) involved ≥3 different stakeholder groups, with the range of these shown in Multimedia Appendix 2. Family carers were the most frequently involved group (27/31, 87% of the studies), followed by people living with dementia (23/31, 74%). Most papers gave little detail about recruitment methods or the background of their PPI representatives. Where papers mentioned attempts to recruit diverse viewpoints, this generally referred to involving different stakeholder perspectives (eg, patients as well as carers and health care professionals), and where inclusivity was dwelled on, this usually referred to the involvement of people living with dementia. Some papers mentioned the linguistic mix or geographical spread of those involved, in particular [52,76,77]. Discussions of racial or ethnic diversity within PPI groups were almost entirely absent. There was one exception [79] in which the inclusion of First Nations representatives was central to the study.

**When Were They Involved?**

The NIHR recommends PPI throughout the research cycle, highlighting in particular 4 key areas where PPI can take place [21]. The studies in this review were mapped to these stages, as shown in Table 2. In total, 29% (9/31) of the studies had involvement from patient and public partners at multiple stages throughout the research cycle [17,49-53,56,57,66], although
sometimes a lack of detail on methods meant that this multistage involvement was presumed (eg, description of a steering group providing oversight “at key milestones” [66] without further description).

As shown in Table 2, a total of 68% (21/31) of the sources did not fit into this NIHR description. These were papers describing a co-design or participatory design process in which the stakeholders involved were both the “co-researchers” or “co-designers” and yet were also positioned as research participants. Typically, these studies involved stakeholders in the iterative design process of a technology prototype. Participants collaborated with the research team on the design process while also being positioned as study participants, for example, being observed testing prototypes or providing feedback as part of interviews. Their involvement could not clearly be classified as designing the research (the study protocol having been designed before their input) or quite as “undertaking/managing the research” as they were the targets of data collection, not involved in the process of collecting them themselves. However, as the authors positioned these stakeholders as collaborators or co-designers along with the research team, these studies were not excluded in the same way that others were when they were more clearly set up as qualitative studies (eg, a focus group to collect end users’ views on technology where ethics approval had been sought for this research process).

How Were They Involved?

Approaches to PPI varied (Table 1). Every paper mentioned at least one form of group activity for their involvement work. Varying terms were used for this—co-design workshops and focus groups were the most frequently mentioned, along with group meetings, group discussions, prototype testing sessions, consultation rounds, group consultations, informal meetings, group feedback sessions, and workshops. Most papers (26/31, 84%) described more than 1 type of activity. In addition to group activities, many conducted interviews [37,49,51,52,55,56,65,66,81,82], observations [54,57,62,65,81,82], or surveys or questionnaires [37,54,65,80,82,83]. A total of 19% (6/31) of the studies set up steering groups that were regularly involved in the research process, described variously as a research development group [53], living experience expert reference group [17], steering group [51,66,71], and patient advisory board [76,77], although details were minimal or absent about what this entailed. A total of 19% (6/31) of the papers [17,49-51,57,74] described some form of integration of patient and public partners (or, in 1 case, a PPI officer as proxy for the PPI group itself [74]) into the main research team, although, again, details were often very minimal about what this entailed. In total, 13% (4/31) of the papers were coauthored by patient and public partners [17,49-51].

**Barriers to and Facilitators of Effective PPI**

A total of 74% (23/31) of the papers included at least some reporting of either facilitators or barriers faced when conducting PPI. In many cases, this reporting was minimal, for example, listing one challenge the team faced. Only 26% (8/31) of the papers had what we considered to be a more thorough discussion of barriers or facilitators [51,56,60,63,64,69,70,73,74,78] (the papers by Hendriks et al [63,64] refer to the same study, as do those by Kowe et al [69,70]). Barriers and facilitators were grouped into themes (Textboxes 4 and 5). Facilitators often focused on ways to achieve richer, more meaningful involvement, for example, working with multiple stakeholder groups and creating a trusting, supportive group dynamic. The barriers identified principally regarded issues with working with dementia as a condition as well as practical issues such as time and budget.
Textbox 4. Facilitating effective patient and public involvement (PPI).

A person-centered approach: choices and adaptability in involvement

1. Prioritizing the well-being and positive experience of those involved [60,77]
2. Offering choices and being led by those involved (how to take part, methods, environment, and level of involvement) [17,69,82]
3. Use of extra time and flexibility for people with dementia, including modification of activities to make them more accessible, acknowledging that there will be no one-size-fits-all [60,77]

Building the group: rapport, trust, and equality

1. Spending time developing group relationships, finding commonalities, and building connections within the team [51,73,77]
2. Building time for chatting and eating together. Informality helps build rapport and flatten hierarchies [60,63,65]
3. Use of a nonresearcher as facilitator [69]
4. Being face-to-face rather than web-based [77]
5. Use of small groups [60]

Multiple viewpoints

1. Including views from multiple stakeholder groups as a way of improving the quality of involvement work and the richness of the data gathered [78,82]
2. Planning a range of methods to recruit and work with different groups (carers, people with dementia, health care professionals, and others) seen as important [82]
3. Considering ways of involving seldom heard groups—from practical adaptations (researchers traveling and not asking patient and public partners to do so) [80] to cultural considerations [79]
4. Group members from different backgrounds learning from each other [65]
5. Support to access different groups was beneficial (eg, working with community organizations or having managerial support to enable staff to take time away from their main role [56])

The right environment

1. Considering accessibility and proximity to local amenities and transport [60,78]
2. Considering who owns the environment—researchers going to those involved (eg, integrating into nursing home environment) may help create a greater sense of equality, flatten hierarchies, and support researchers’ understanding of the group they are working with [56,63,80]
3. Being face-to-face enabled hands-on workshops, improved group dynamics, and reduced technology barriers [77,78]
4. Web-based environments enable geographically diverse groups to come together and may keep discussions more focused [73]

Support and training

1. Having facilitators or members of the research team who are skilled and experienced working with people with dementia [60,78]
2. Providing training for patient and public partners (eg, data collection, thematic analysis, and computer skills) [51,60]
3. Supporting patient and public partners with adequate time to reflect and debrief with members of the academic research team [51]
4. Using paper prototypes to overcome technology barriers [66]
5. Providing adequate support for people with dementia (family carers [60] or modified activities [60,77,78])

Pragmatism and compromise

1. Proxy involvement (of family, PPI officers, and nursing home staff) used in place of people with dementia (or people with moderate to severe dementia) in cases in which their involvement was not seen as feasible [55,61,74]
2. One-to-one sessions found to be easier to organize than group sessions [53]
3. Virtual meetings may be easier to organize than face-to-face meetings [73]
Textbox 5. Barriers to effective patient and public involvement (PPI).

The nature of dementia
1. Cognitive impairments seen as too great a barrier to attempt PPI with people with dementia [55,61]
2. Input from people with dementia described as very minimal [62]
3. Attempts at adaptations unsuccessful [63,64]
4. Variations in presentation making it difficult to plan a particular approach or manage a group [63,64]
5. Carers and people with dementia both overestimating the abilities of the latter [63,64]
6. Unreliable historians— for example the challenge of interpreting someone’s account of their ability to participate in activities of daily living while they also recount recent interactions with long-dead relatives [63,64]
7. The emotional load faced by researchers working with this group, including challenges such as being asked for support or advice out of their scope [63,64]

Inequality of relationships within the group
1. Some authors highlighted issues with patient and public partners feeling undervalued or not equal within the team; this applied to those without dementia [51,63], though dementia was also seen as an additional barrier to a sense of equality [63,64]
2. Lack of payment for PPI also contributed to this, as well as the limited scope or lack of defined roles and responsibilities for patient and public partners [63,69]

Time pressures
1. Researchers’ time pressures—co-design or other involvement activities as time-consuming processes that can be difficult to manage alongside the time pressures of a research study [75,81,83]
2. Family carers’ time pressures—busy schedules and stressful lives [37,69]
3. Staff time pressures—nursing home staff and health care professionals’ strict shift patterns and limited flexibility for time away from work [80], in some cases exacerbated by the COVID-19 pandemic [52]
4. Limited time resulting in reduced or inadequate training for patient and public partners [51,69]
5. Rapport building in the group suffering as a result of lack of time [51,56]

Recruitment and diversity
1. Small numbers of people involved resulting in reduced diversity of opinions and a poorer representation of stakeholders [56,57,75]
2. Challenges with generating interest in the study or reaching particular groups [74,78] and COVID-19 causing staffing pressures [52] and a lack of face-to-face options for people with dementia [82]

Processes and communication
1. Communicating complex content (the ethics of smart homes, technology use, and design processes) was particularly challenging for people with dementia [62,68]
2. Use of jargon terminology by researchers was a barrier for all patient and public partners (not only those with dementia) [63,69]
3. Methods of communication—use of phone for people with dementia was limiting [82], and sending too many emails was unpopular [56]
4. Processes for PPI members experienced as boring or repetitive (eg, completing multiple assessments), especially when combined with a lack of communication about the purpose or the results of their input [56]
5. Lack of involvement and communication early in the study resulting in stakeholders having less of a connection or understanding of the project or feeling that their contributions were less valued [51,63]

Impact of PPI
Although most papers (28/31, 90%) implied or briefly commented that stakeholder involvement had some impact on their study (usually on the iterative design process), this was sometimes without any description of what the impact was. Where any details were given, as was the case in 52% (16/31) of the papers [51,54-57,60,61,63,69,73-79], the results are summarized in Textbox 6.
Impact on the research

- PPI activities helped set groups’ research agendas, with clearly defined stakeholder priorities for research [76,77]. PPI data were identified as something that can be shared with and used by the wider research community when planning research [77].
- Involvement in research design resulted in a set of recommendations that the authors hope will improve the acceptability for research participants [78] and in specific cultural adaptations and approaches [79].
- Involvement in data collection was reported as adding richness to the data on account of the rapport and connections that peer researchers built with the people with dementia they were interviewing [51].
- Many papers (21/31, 68%) commented that the methods used (eg, co-design and participatory design) had an end result that was in some way grounded in the views or priorities of users but often with minimal detail. In total, 13% (4/31) of the papers [55,57,74,75] gave detail about the extent to which user groups drove the development or design of technology, reflecting on the value of their input.
- Coresearcher involvement in a thematic analysis workshop made for a more robust analysis, with differing perspectives between the research team and coresearchers showing the need for more PPI at the analysis stage in the future [69]. The limited impact that PPI activities had at the analysis stage was also reflected on, citing inadequate time and training for coresearchers resulting in brief and surface-level group analysis sessions [51,69].

Impact on patient and public partners

- Feedback on positive experiences of patient and public partners was provided in general terms [74,78]. Positive relationships between team members were developed, with feelings of mutual respect as well as the value of finding connections being reported [51,56,73].
- Some reported empowerment and satisfaction with the project and their role in it [54,60,61].
- Patient and public partners developed new skills [51].
- Negative experiences were reported on, including finding tasks boring or repetitive or processes complex [56]. Some papers also reported that patient and public partners felt underinvolved [51] or not treated as equal partners [63].

Impact on the academic research team

- Researchers gained a deeper understanding of the needs and priorities of the group they were seeking to conduct research with and for [79].
- Researchers developed a sense of connection with and respect for other disciplines or stakeholders they had not previously worked with [60,65].
- One paper reflected on the emotional burden associated with close working with people with dementia and the need for support for researchers as well as the people with dementia themselves [64].
- The initial challenges of stepping back when sharing responsibility with peer researchers was reported on, which became easier with experience [51].

In general, no formal methods were used for evaluating the impact of PPI activities. Where papers reported on impact, it was usually limited to the authors’ reflections, including when reporting on the impact on patient and public partners. In the case of 13% (4/31) of the papers [51,56,63,73], the authors reported seeking direct feedback from those who had been involved, for example, in the form of interviews; evaluation forms; or, in 3% (1/31) of the studies, a much more extensive retrospective analysis using formalized methods [63].

Discussion

Principal Findings

In this scoping review, we set out to explore the concept of PPI in technology-related dementia research. The papers reviewed in this study revealed that dementia researchers are embracing PPI, with varied and sometimes ambitious methods, values centered on inclusivity and coproduction, and involvement of a range of stakeholder groups. We found that approaches often blurred boundaries between those involved as “researchers” and those involved as “participants” so that most studies’ (21/31, 68%) PPI activities did not fit into a strict definition of PPI, for example, as set out by the NIHR [32]. Although the involvement activities being undertaken demonstrate this to be a rapidly expanding and developing field, the brevity in the reporting of such activities (often without comments on the impact of PPI) perhaps highlights the need for clearer reporting guidelines. Where mentioned, the impact of PPI was generally reported as being positive on research quality, patient and public experience, and the learning experiences provided to researchers. We comment further on our objectives in the following sections.

Objective 1: To Describe the Approaches to PPI Used to Date in Technology-Related Dementia Research (Exploring Who Is Involved, When, and How)

We found that there was a narrative across many of the included papers about the value of involvement and coproduction methods, with many authors describing their aspiration for high levels of involvement with a sense of partnership and equality with stakeholders. A few consciously excluded people with dementia from this aspiration, citing cognitive impairments as making it either practically or ethically too challenging to involve this group. These views were chiefly expressed in older papers (before 2010). More recent papers were broadly inclusive, with some describing their efforts to involve people with dementia along with other stakeholder groups such as...
carers, health care professionals, and older adults without dementia. Sometimes, these groups were involved in similar ways, and sometimes, there were 2 very separate approaches, for example, a set of workshops with people with dementia and carers and more extensive involvement of health care professionals or others without dementia in the research process (eg, playing a role in designing the protocol as members of a steering group). The fact that a significant majority (26/31, 84%) involved more than one stakeholder group, with many involving ≥3 groups (17/31, 55%), reflects the value placed on hearing from multiple viewpoints.

Despite this widespread acknowledgment of the value of collaborative or coproduction methods, it was not always clear from the papers to what extent their methods reflected these values. Some used methods that perhaps lend themselves better to a consultative approach (such as one-off focus groups or surveys) rather than a collaboration or coproduction approach [20]. Consultative methods have some value in enabling researchers to find out more about people’s views and experiences. They are also relatively easy to organize (often one-off events as opposed to longer-term involvement), meaning that they are a practical way of hearing from a wide range of stakeholders [20]. However, these methods mean that involvement will always be more limited in scope—there is no 2-way discussion, and there is the risk of disengagement from stakeholders who feel that their views are not being listened to [20], as indeed was reported by Hendriks et al [63]. Where stated methods or approaches may have enabled more of a partnership approach to take place (eg, the use of a series of co-design workshops or involvement of a steering group), the brevity of the write-up often prevented us from understanding whether or how this happened.

The most common methods of involvement were through some form of group activity such as focus groups, workshops, prototype testing sessions, or roundtable discussions. Every paper mentioned some form of group activity such as these, with or without other methods. Involvement often occurred at just one stage of research—most commonly, this was contributing to the design of a prototype but with no involvement either before or after this (eg, in protocol design, recruitment, analysis, or dissemination), although there were exceptions to this, with studies involving patients or the public throughout multiple stages of the research.

Most studies (26/31, 84%) used multiple methods of involvement. There will be practical reasons for using different methods, some suiting particular groups or settings more than others, especially when we consider involving people with dementia and the flexibility of approach that this requires. Tiersen et al [82] described many different methods in their paper, reflecting that this “resulted in triangulation of investigators, methods and data sources to develop a more comprehensive understanding of the phenomena being studied.”

The use of multiple methods also allowed for more flexibility of involvement, with some able to take on a more active or sustained role than others as able or desired. This was cited as an aspiration or suggestion for future PPI by groups who did not have such flexible approaches, such as Kowe et al [69]. The paper by Liddle et al [17] described the flexibility of roles and high levels of involvement. Their “Living Experience Expert Reference Group” involved 15 people with dementia and carers, of whom 4 (2 people with dementia and 2 carers) were also integrated into the main investigator team. Roles included helping set research priorities, developing the interview topic guide, analysis of data, reflexivity sessions, and the write-up of the study. Shadarevian et al [49] and Hung et al [50] also described the integration of stakeholders into the main research group, mentioning roles in undertaking and managing the research along with analysis and dissemination, although there was little detail about the methods used for this involvement. PPI in data collection was rare. It was described in most depth by Daly Lynn et al [51], who worked with “peer researchers” (older adults without dementia) to interview research participants with dementia about their experiences with smart home living, with an insightful write-up detailing both the positive experiences and the challenges they faced.

Finally, although the notion of diverse viewpoints being included in PPI was often celebrated, this almost always referred to the inclusion of a variety of stakeholder groups (eg, patients, carers, and health care professionals). In general, there was very little explicit discussion of the demographics of PPI representatives. To reduce health-related inequalities and bias, researchers should consider not only how to involve people with dementia but also the demographics of this heterogeneous population, explicitly seeking ways to involve underrepresented groups.

**Objective 2: To Describe the Reported Barriers to and Facilitators of Effective PPI in This Area**

The general lack of evaluation or reflections on barriers to and facilitators of involvement means that the themes described in this paper result from a minority of studies, with most derived from just 26% (8/31) of the studies [51,56,60,63,69,73,74,78]. Themes drawn out broadly matched those frequently documented in the literature [11,25,41,42], in particular barriers such as time and budget, recruitment issues, and the specific challenges of adapting activities to be suitable for people with dementia.

Facilitators tended to focus on the manner in which activities were carried out (eg, how informality helped flatten hierarchies) rather than on specific methods or approaches, such as focus groups or interviews. The themes here mirror the key principles of coproduction outlined by the NIHR [21,32]. The NIHR emphasizes that coproduction does not require a specific method but a more nuanced focus on interpersonal skills, relationship building, and power sharing (Textbox 1) [32]. However, although many papers stated the need for approaches using these principles or claimed to have worked with such values in mind, details about what was done were often limited. We would welcome further and more detailed reporting on these activities so as to build the knowledge base among research teams and enable more high-quality PPI to be conducted in the future with this population. As also emphasized by Hendriks et al [64], there is a need for more than anecdotal evidence in the literature about how to involve people with dementia—the lack of guidelines or a strong evidence base makes progression challenging.
The barriers identified reflect the challenging nature of PPI in technology-related dementia research. Time and resources were frequently cited as limiting factors. Some studies avoided the challenge of making PPI activities accessible to people with dementia by not involving them at all. Hendriks et al [64] detailed the challenges of involving people with dementia in a particularly frank manner. They reported on the difficulty of sufficiently modifying activities to make them accessible, difficulties with variability in dementia presentations, and overestimation of abilities by people with dementia and their family carers. They went as far as to say that “the differences between the designer and the person to design for are too big to speak about equality in participation” [64]. A few papers highlighted the complex topics under discussion (eg, smart home technology, ethics, and design processes) and the difficulty of translating these issues into something someone with dementia could understand and contribute to [62,63,68]. Kort et al [62] went as far as to say that the complexity of content meant that input from people with dementia was very basic, commenting that “the actual participation in the project was deemed more important than the actual contribution.” Although pessimistic sounding in tone at times, these honest admissions of the challenges faced and the inability of researchers to overcome them to a meaningful degree help the research community understand the current landscape of PPI in technology-related dementia research. They reinforce the need for significant investment to be made for the involvement of people with dementia to be successful. This is not merely in the creation of accessible resources or the provision of the right environment. Rather, it may be that significant cultural shifts need to take place for researchers to be able to plan and conduct effective coproduction based on the key principles of power sharing, inclusivity, respect, reciprocity, and relationship building [32]. It seems likely that researchers would benefit from significant training and support in understanding power dynamics and coproduction as well as support to learn more about how to work with people with cognitive impairments.

**Objective 3: To Examine and Report on the Impact of PPI in This Area**

Papers with more than a brief comment on the impact of involvement were scarce in this review, mirroring the findings by Suijkerbuijk et al [41] as well as the findings of those studies considering PPI in dementia more broadly [25,26,42,84]. A few papers, notably those by Daly Lynn et al [51], Muñoz et al [56], Banbury et al [73], and Hendriks et al [63,64], provided valuable discussions and evaluations of impact.

Where papers reported on the impact of their involvement work, the results were largely positive. The studies demonstrated that PPI can have a positive impact on research quality at multiple stages of the research cycle as well as on those taking part. To achieve this, the authors reflected on the need to involve multiple stakeholder groups and use multiple methods of involvement to provide a person-centered and flexible approach in which people feel well supported and valued for their contribution. To do this evidently requires significant investment of time and resources. Even those papers detailing what we considered to be relatively high levels of involvement spoke about the need for more time and resources, for example, to improve levels of training or offer more formalized PPI roles [51,69].

It was helpful to see negative experiences with PPI also reported [56,63,64]. These are often missing from the literature [85] but provide helpful learning points. Another area that received little attention is the emotional impact of PPI on researchers. This was commented on briefly by Hendriks et al [64] but otherwise did not feature in the studies we reviewed, although it has been noted as a feature of PPI in dementia research more broadly [43,84]. The emotional impact on researchers might be seen as both positive and negative—as a research community, it would be helpful to recognize and value the learning and increased empathy that can come from close working with patient and public partners. It is also important to recognize the potential distress or emotional burden felt by researchers, which might be associated with this relationship, in particular for those unused to working with people in cognitive decline [43]. It is important that future studies consider these impacts and that researchers as well as patient and public partners have access to adequate training and support.

Across most studies (27/31, 87%), formal or standardized methods were not used to capture impact. Reporting was generally limited to the authors’ personal reflections. In only 13% (4/31) of the papers did the authors report seeking direct feedback from those who had been involved, for example, in the form of interviews or evaluation forms [51,56,63,73]. Hendriks et al [63] included a detailed evaluation of impact. This team retrospectively analyzed their participatory design process, mapping out the decision-making process at each stage of the project and considering the extent to which coresearchers had been involved and, therefore, whether participation had been truly meaningful. They also interviewed some of those who had been involved and analyzed themes that emerged. There is a risk that reporting on the impact of PPI can lead to an overvaluing of that which is easily measured but of little meaning (such as the number of people involved) instead of these more complex issues such as research culture or power relations [85]. Therefore, this example by Hendriks et al [63] reflecting on and evaluating the power dynamics at play and the processes that took place is particularly commendable.

In the future, it would be helpful for all involved—researchers, patients, and the public—if there were more recording and reporting of the impact of involvement to help all parties understand if, when, how, and why partnership working is beneficial [20,43]. Capturing this in a meaningful way is the challenge ahead of us [85,86]. A focus on the dialogue and the learning is felt to be helpful—Russell et al [85] recommend exploring “the complexity and richness of this relationship, using methods that emphasise illumination rather than measurement, and asking when, why, and with whom the dialogue happens or fails to happen.”

**Defining PPI: Challenges We Faced in This Review**

One of the challenges we faced was the varying terminology and approaches used to involve groups in research. We applied a broad definition of “involvement” and, therefore, included papers using co-design or participatory design processes, as described, for example, by Hendriks et al [63]. Within a
participatory design approach, there is a deliberate blurring of the roles of “designer” and “end user.” When written up in a research context, this can lead to a blurring of the roles of researcher, designer, end user, and research participant. From a PPI perspective, it is not usually considered appropriate for people involved in research to also be research participants as this can compromise both the researcher and person involved [20]. However, the NIHR gives the example of participatory or action research as a possible exception to this rule, and it was often these types of studies that we reviewed. Nevertheless, it was often difficult to determine which studies met our inclusion criteria, in part because of this mixing of roles and because of lack of detail in the methodology sections. Where studies explicitly used qualitative research methods (stating a qualitative approach and collecting data for analysis with appropriate ethics approval), they were excluded from this review. However, details were often missing, or sometimes subsections of a study appeared to be qualitative, whereas other sections were framed more like PPI activities. Terminology could not be relied on as it was applied inconsistently among studies. Similarly, information about ethics applications was not always available, and we did not use this as part of the inclusion and exclusion criteria. Had we used a stricter definition of PPI, we would have excluded a significant portion of the literature (21/31, 68% of the studies in this review). We felt that doing this would result in a misrepresentation of the type of involvement work being carried out and in missing key learning points from these studies.

Strengths
We used a comprehensive search strategy considering all types of technology, all types of dementia, and many terms for “patient and public involvement” to reflect the different types of involvement in the field, building on search strategies from other reviews [11,19,25,41,42] that at times had been narrower in scope (eg, looking at “patient and public involvement” but not “co-design”). The 2 reviewers overcame the difficulty of defining PPI through regular communication and close working throughout the screening stages, consulting with a third member of the team where required. The review is further strengthened by multidisciplinary team input, with representatives from health care and health sciences as well as from design and technology backgrounds, which we hope ensures that our reflections and conclusions are of interest and applicable to a wide range of disciplines.

Limitations
We did not conduct a gray literature search. Doing so might have resulted in a broader range of accounts of involvement being included. In addition, we did not involve patients or the public in this review, which may have contributed additional perspective and depth. However, we have planned and started recruitment for a much broader PPI strategy for our research center. The results of this review will be shared with our steering group so as to jointly consider how the findings should inform our PPI work as a center.

Conclusions
At present, most involvement in technology-related dementia research is limited in breadth (often to just 1 stage in the research cycle) and depth (often consultative rather than with any sharing of power). We see across the literature shared aspirations of high levels of meaningful involvement in research, and it is encouraging to see some evidence of this being put into practice, with some reporting on methods used for involvement and the impact this has. Where papers gave details, it appears that a flexible approach with multiple methods used at different stages of the research cycle may be the most successful, tailoring methods to the various groups or individuals involved and facilitating greater depth or breadth of involvement according to people’s wishes and abilities. When this is done well, PPI can have a positive impact on both the research and those involved. This evidently will take significant time and resources, particularly if the approaches used are to move beyond consultations to collaboration or coproduction. Wider reporting of methods and facilitative strategies along with more formalized methods for recording and reporting on meaningful impact would be helpful so that all those involved—researchers, patients, and other stakeholders—can understand and learn how best to jointly conduct research.

Acknowledgments
PK is supported by the National Institute for Health and Care Research Imperial Biomedical Research Centre and Imperial Health Charity. HL is funded by the UK Dementia Research Institute and Imperial College London President’s PhD Scholarship.

Authors’ Contributions
PK wrote the manuscript. PK and HL developed the protocol with supervision from CMA and SD. PK and HL screened the results and extracted the data with third-author support where necessary from CMA. All authors reviewed and contributed to the final manuscript.

Conflicts of Interest
None declared.

Multimedia Appendix 1
Detailed search strings used for the different databases. [DOCX File, 22 KB - aging_v7i1e48292_app1.docx ]
Multimedia Appendix 2
Stakeholder groups included as part of patient and public involvement activities.

[PNG File, 76 KB - aging_v7i1e48292_app2.png]

Multimedia Appendix 3
PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) checklist.

[DOCX File, 83 KB - aging_v7i1e48292_app3.docx]

References


33. Ideo design thinking homepage. Ideo Design Thinking. URL: https://designthinking.ideo.com/ [accessed 2023-04-04]


48. Covidie systematic review software. Veritas Health Innovation. URL: https://www.covidence.org/ [accessed 2022-12-12]


75. Rai HK, Schneider J, Orrell M. An individual cognitive stimulation therapy app for people with dementia: development and usability study of thinkability. JMIR Aging 2020 Nov 16;3(2):e17105 [FREE Full Text] [doi: 10.2196/17105] [Medline: 33916451]


Abbreviations

NIHR: National Institute for Health and Care Research
PPI: patient and public involvement
PRISMA-ScR: Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

Edited by Y Jiang; submitted 18.04.23; peer-reviewed by O Klein, A Sharma; comments to author 20.07.23; revised version received 12.09.23; accepted 17.01.24; published 04.03.24.

Please cite as:
Patient and Public Involvement in Technology-Related Dementia Research: Scoping Review
JMIR Aging 2024;7:e48292
URL: https://aging.jmir.org/2024/1/e48292
doi: 10.2196/48292
PMID:38437014

©Pippa Kirby, Helen Lai, Sophie Horrocks, Matthew Harrison, Danielle Wilson, Sarah Daniels, Rafael A Calvo, David J Sharp, Caroline M Alexander. Originally published in JMIR Aging (https://aging.jmir.org), 04.03.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Review

Strategies to Mitigate Age-Related Bias in Machine Learning: Scoping Review

Charlene Chu1,2,3,4*, BScN, GNC(c), PhD; Simon Donato-Woodger1*, BScN; Shehroz S Khan2,5*, PhD; Tianyu Shi1,6*, MSc; Kathleen Leslie7, BScN, JD, PhD; Samira Abbaspaghilizadeh-Rahimi8*, BEng, PhD; Rune Nyrup9*, PhD; Amanda Grenier10*, PhD

1Lawrence Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada
2Knowledge, Innovation, Talent, Everywhere (KITE), Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada
3Institute for Life Course and Aging, Faculty of Social Work, University of Toronto, Toronto, ON, Canada
4Rehabilitation Sciences Institute, University of Toronto, Toronto, ON, Canada
5Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada
6Department of Civil Engineering, University of Toronto, Toronto, ON, Canada
7Faculty of Health Disciplines, Athabasca University, Athabasca, AB, Canada
8Department of Family Medicine, McGill University, Montreal, QC, Canada
9Centre for Science Studies, Department of Mathematics, Aarhus University, Aarhus, Denmark
10Factor-Inwentash Faculty of Social Work, University of Toronto and Baycrest Hospital, Toronto, ON, Canada
*these authors contributed equally

Corresponding Author:
Charlene Chu, BScN, GNC(c), PhD
Lawrence Bloomberg Faculty of Nursing
University of Toronto
155 College Street
Toronto, ON, M5T 1P8
Canada
Phone: 1 416 946 0217
Email: charlene.chu@utoronto.ca

Abstract

Background: Research suggests that digital ageism, that is, age-related bias, is present in the development and deployment of machine learning (ML) models. Despite the recognition of the importance of this problem, there is a lack of research that specifically examines the strategies used to mitigate age-related bias in ML models and the effectiveness of these strategies.

Objective: To address this gap, we conducted a scoping review of mitigation strategies to reduce age-related bias in ML.

Methods: We followed a scoping review methodology framework developed by Arksey and O’Malley. The search was developed in conjunction with an information specialist and conducted in 6 electronic databases (IEEE Xplore, Scopus, Web of Science, CINAHL, EMBASE, and the ACM digital library), as well as 2 additional gray literature databases (OpenGrey and Grey Literature Report).

Results: We identified 8 publications that attempted to mitigate age-related bias in ML approaches. Age-related bias was introduced primarily due to a lack of representation of older adults in the data. Efforts to mitigate bias were categorized into one of three approaches: (1) creating a more balanced data set, (2) augmenting and supplementing their data, and (3) modifying the algorithm directly to achieve a more balanced result.

Conclusions: Identifying and mitigating related biases in ML models is critical to fostering fairness, equity, inclusion, and social benefits. Our analysis underscores the ongoing need for rigorous research and the development of effective mitigation approaches to address digital ageism, ensuring that ML systems are used in a way that upholds the interests of all individuals.

Trial Registration: Open Science Framework AMG5P; https://osf.io/amg5p

(JMIR Aging 2024;7:e53564) doi:10.2196/53564

https://aging.jmir.org/2024/1/e53564 Jmir Aging 2024 | vol. 7 | e53564 | p.792 (page number not for citation purposes)
Introduction

The rapid progress of machine learning (ML) has revolutionized health care decision-making, medical diagnosis, and other domains [1]. However, as the influence of ML systems expands, so do concerns regarding potential fairness issues that may arise from ML systems encoding human biases [2]. As an example, population health management systems have been found to underestimate the health risks facing Black patients, who are typically underrepresented in health care data due to systemic challenges accessing health care [3]. Concurrent with the rise of ML, there has also been a growing demand for efforts to improve the fairness of ML systems by better representing systemically disadvantaged groups in their data [4,5], such as gender and ethnic minority individuals [6-8]. To better understand the scope of this problem, frameworks to classify the various forms of bias present in ML have been developed. Our previous work used the framework developed by Mehrabi et al [5], which classified numerous types of bias according to the characteristics of each bias as well as where it would be introduced into an ML system in the cycle of providing training data (data to algorithm), the ML model interacting with the public (algorithm to user), and the public’s data being used for future testing (user to data). In our earlier investigation, we delineated 9 distinct categories of ML bias that could provide avenues for age-related bias to affect ML systems [9], using the conceptual framework by Mehrabi et al [5]. In this investigation, the prevalent forms of bias from their framework were: (1) representation bias, which emerges when the data set used for training the ML model inadequately reflects the diversity of the user population compromising the performance for specific demographic groups; (2) evaluation bias, which can occur when the model is tested with unrepresentative data and inappropriate evaluation benchmarks are used; (3) aggregation bias, in which distinct demographics within a larger sample are categorized in a way that makes their unique characteristics indistinguishable; (4) algorithmic bias, where the algorithm itself is the origin of the bias leading to distorted outcomes; and (5) measurement bias, which arises from how certain features are selected, measured, and used. When data are measured or gathered using improper tools or techniques, the resulting evaluation of the data by an ML model does not reflect the relevant variables within the data [5,9].

Digital ageism is a form of ageism perpetuated through the development, use, and deployment of technology and ML models [10,11]. Recently, the World Health Organization released a brief report about age-related bias in ML models and raised critical questions about the equitable treatment of older people across various sectors [12]. The rising concerns about digital ageism highlight the pressing need for further research and policy interventions to address the potential biases and discriminatory practices that may affect older adults in the digital era [13-15]. Recent studies have demonstrated instances of digital ageism, emphasizing the urgency of designing and deploying technologically inclusive solutions to ensure equitable treatment and opportunities for individuals of all ages [16]. The exclusion of older adults from the development of digital technologies has been previously researched [15,17,18] and can manifest in many ways. Older adults may not be adequately represented in training and testing data for ML models, resulting in models with reduced accuracy for older adult data and being vulnerable to multiple intersecting disadvantages [19]. For example, older adults who live in long-term care homes may have limited access to the internet and may be excluded from technological advances [20]. Data may also aggregate older adults into arbitrary age blocks, replicating problematic assumptions that link functional decline with age and failing to represent the diversity of the older adult population [10]. The marketing strategies for these ML systems often highlight their use in health care, reinforcing the idea of aging as a period of physical and mental decline [21]. As ML models and technologies become inextricably part of accessing opportunities and services, older adults’ risk of being left behind by a growing digital divide increases [22]. This is particularly alarming considering that the older population represents the fastest-growing demographic worldwide [23].

The topic of digital ageism is gaining prominence in scholarly discussions, leading researchers to investigate these phenomena from various perspectives [9,19,24,25]. Previous investigations have focused on developing conceptual frameworks to comprehend and define the nature and implications of digital ageism [13]. Previous reviews of facial image data sets have also found that older adults, particularly older adults aged 85 years or older, are underrepresented in a majority of data sets [18]. While this research has been foundational in identifying and characterizing these biases, there is now a critical need to focus on the mitigation strategies that can address age-related bias in ML models. The purpose of this scoping review is to advance this crucial discussion by shedding light on the mitigation strategies currently being used to address age-related bias in ML models. By bridging the gap between theory and practice, this research aims to pave the way for meaningful and impactful interventions that can rectify biases and promote inclusivity in the digital age. Our research focuses on two main questions: (1) Which mitigation strategies have been used to address age-related bias in artificial intelligence, and how successful were these strategies? and (2) Specifically, what types of biases were targeted and mitigated during these efforts?

Methods

Overview

This review is part of a larger scoping review about digital ageism that follows a 6-stage methodology framework developed by Arksey and O’Malley [26] and further refined by Levac et al [27]. A scoping review was appropriate to explore...
the study’s aims of summarizing the available evidence, identifying gaps, and establishing future directions regarding mitigation strategies. As the breadth and depth of the literature are unclear, a wide and interdisciplinary approach was used [10]. The description of the review was published elsewhere [28] and registered in the Open Science Framework database [29]. This review also followed the PRISMA-ScR (Preferred Reporting for Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews) format [30].

Information Sources and Search
An information specialist helped develop the search strategy in Scopus, which was then translated into 5 other databases (Web of Science, CINAHL, EMBASE, IEEE Xplore, and the ACM digital library). The search strategy included the terms “machine learning,” “artificial intelligence,” “algorithms,” “neural networks,” “deep learning,” “algorithmic bias,” “biased,” “discrimination,” “ageism,” “age,” and “older people.”

Eligibility Criteria
Articles were included if they were published in English and focused on “artificial intelligence” in the context of algorithms that make predictions and classifications about data; “bias”; and age-related terms such as “aging,” “older,” and “demographic.” As the term “artificial intelligence” is over 50 years old [31], the search strategy was also not restricted by publication date, and databases were searched from inception. Papers were excluded if they included nonhuman topics. Theses, conference abstracts, dissertations, nonpeer-reviewed conference proceedings, books and book chapters, perspectives, editorials, and editorial letters were also excluded.

Selection of Sources of Evidence and Charting the Data
The academic literature search was completed in January 2022. All citations were uploaded to Covidence (Veritas Health Innovation), a systematic review software, and duplicates were removed. The titles and abstracts of all articles were screened by 2 independent reviewers according to the eligibility criteria. Once the abstract screening was complete, the full text of each article was reviewed by 2 independent reviewers to judge the article’s relevance to the research questions. Data extraction included the manuscript information (title, authors, year, and location), study design, type of ML model and purpose, database used, type of data, presence of age-related bias, mitigation strategy used, and effectiveness of the strategy if reported. The framework by Mehrabi et al [5] (Figure 1 [5]), which identified different sources of biases that can affect ML according to the data-to-algorithm (data), algorithm-to-user (modeling), and user-to-data (deployment) interaction loops, was used in this review to identify the different types of bias in the included studies [5]. A total of 5 of the 19 different types of biases in the framework by Mehrabi et al [5] (Figure 1) were included in the extraction table, including representation and evaluation bias, aggregation bias, measurement bias, and algorithmic bias.

Studies were selected if they acknowledged the presence of any bias against older adults in either their data or results, and the researchers then took any action to correct that bias, regardless of its effectiveness. For example, publications were selected based on whether authors attempted to enhance the performance of their model on older demographics, regardless of the success of their efforts. Biweekly meetings were held to discuss the progress of the charting process. Disagreements were resolved through discussion or by having the first author (CC) act as a third reviewer. The extracted information was converted into table format, which allowed the authors to develop a narrative description according to the type of mitigation strategy (Table S1 in Multimedia Appendix 1 [32-42]). The team conducted additional analysis of the databases in the included studies to identify data disparities among older adults and provide further directions for future studies in the field of digital ageism. The results of the literature search are reported in tables. One of the challenges involved in assessing the inclusion of older adults involves defining the age at which someone is considered “old.” While the commonly accepted age for legal recognition as a “senior citizen,” based on general eligibility for a public pension, is 65 years [43], the data sets and articles we reviewed grouped older adults into a much wider range of age categories (Table S2 in Multimedia Appendix 1), starting with adults aged 50 or older. When we refer to “older adults” throughout this paper, we are referring to either (1) the participants in the data set aged 60 years or older, or else (2) the oldest age category found in the data set or publication being discussed.
Figure 1. The framework by Mehrabi et al for bias in machine learning. The specific biases discussed in this review are underlined.

1. **Measurement bias**
2. Omitted variable bias
3. **Representation bias**
4. Aggregation bias
5. Sampling bias
6. Longitudinal data fallacy
7. Linking bias

1. **Algorithmic bias**
2. User interaction bias
   a) Presentation bias
   b) Ranking bias
3. Popularity bias
4. Emergent bias
5. **Evaluation bias**

Ethical Considerations

Ethics approval was obtained from the (University of Toronto) Research Ethics Board (REB #40095) for a larger study on the same topic. This study does not contain any studies with human participants performed by any of the authors.

Results

Overview

From our search, 14,611 academic publications were identified. After removing duplicates, we screened the abstracts of the remaining 7903 publications. During the abstract screening process, we excluded 7592 publications. Subsequently, we conducted a full-text screening of 310 articles, ultimately including 8 academic publications in this review (Figure 2).
Types of Ageist Bias in the Selected Publications, per the Framework by Mehrabi et al

After reviewing the full text of each of the 8 publications, 5 types of bias found in the framework by Mehrabi et al [5] (Figure 1) were identified. Representation bias occurs when the data set underrepresents or misrepresents specific demographics within the overall population, resulting in a nonrepresentative data set [5]. Evaluation bias is similar to representation bias: it occurs when inappropriate evaluation benchmark data are selected to assess ML models [5]. In this review, this amounted to using the same underrepresentative data sets for training as well as testing the model, which was found to be the case in 7 publications [32-38]. Aggregation bias occurs when conclusions are drawn based on observations about a larger group, overriding unique characteristics about a smaller demographic within that group [5]. As a result, the data set fails to account for the unique characteristics of more specific demographics within the overall data set. Aggregation bias was found in 2 publications in this review [34,35]. Algorithmic bias occurs when the bias is generated at the level of the algorithm’s calculation itself rather than being a by-product of biased data or measurement tools being provided to that algorithm. Algorithmic bias was found in 2 papers in this review [36,39]. Measurement bias occurs when the data being processed by the algorithm fail to represent the variable of interest accurately. It can often arise from the methods used to collect or measure the data or respective variables. Measurement bias was demonstrated in 1 publication in this review [35]. A complete list of each type of bias found in each publication, along with a rationale, can be found in Table S1 in Multimedia Appendix 1.
Data Sets Used in the Included Studies

Table S2 in Multimedia Appendix 1 presents the demographic breakdown of the data based on age to determine the extent of underrepresentation of older adults in common data sets. Overall, there was a large data disparity between the data of older adults compared with younger individuals in all the databases in this review. In the FG-Net data set, the oldest age group was “aged 61 years old or older.” In this age group, the FG-Net data set only had 7 images (0.7%) [40]. In the MORPH Academic data set, there were only 3933 (7.1%) images of individuals in the “aged between 50 and 77 years” category [44], and in the MORPH Longitudinal data set, only 5615 (1.4%) of the data set’s 402,055 images were from the “aged between 60 and 69 years” (5021 images) and “aged 70 years or older” (594 images) categories [45]. The CACC data set contains 163,446 images, divided into age groups of 10 years (0-10 years, 10-20 years, 20-30 years, 30-40 years, 40-50 years, 50-60 years, and 60 years or older). However, only 2912 (1.78%) of these images depict participants older than 60 years (the fifth figure in Georgopoulos et al [34]). Grouping all older adults into one category of aged 55 years or older increases the likelihood of aggregation bias [5].

The APPA-Real data set contains 7591 images (from public internet repositories), aged between 10 and 95 years. When we combine the APPA-Real data set’s 4 oldest age groups: between 60 and 70 years (254 images), between 70 and 80 years (111 images), between 80 and 90 years (68 images), and between 90 and 95 years (13 images), the combined total of 446 images accounts for just 6% of the entire data set [41]. Lastly, the 100 Celebrities data set created by Jung et al [35] has a smaller-sized balanced data set to offset the imbalances in the IMDB-Wiki and Twitter (subsequently rebranded as X) Profile data sets also used in the same study. Participants were divided into 3 age groups: between 14 and 34 years (33 images), between 35 and 55 years (34 images), and 55 years or older (33 images), with the “55 years or older” age group having equal representation with the other demographics, although grouping all older adults into 1 category of aged 55 years or older increases the likelihood of aggregation bias [5].

For the other data sets present in this review (11,000 Hands, ABIDE, CoRR, DLBS, NKI Rockland, Pilots Parliament Benchmark, and a data set comprised of Twitter Profile data), the exact data for the age demographics in each data set was not readily available (Table S2 in Multimedia Appendix 1).

Bias Mitigation Strategies

We found 8 studies that attempted to mitigate bias against older adults. The studies were all related to “computer vision” systems, systems that rely on an ML model’s processing and interpretation of images, although with varying aims: 1 study focused on hand images [32], another study focused on radiological scan interpretation [36], and 6 others focused on facial images [33-35,37-39]. A complete list of the papers, the data sets that were used, and the strategies used to mitigate bias and their outcomes can be found in Table S1 in Multimedia Appendix 1. We identified 3 broad categories of bias mitigation strategies: data set balancing, data set augmentation, and algorithm alterations. This section will provide a comprehensive overview of each bias mitigation strategy and its effectiveness [34,36].

Data Set Balancing

Data set balancing involves the practice of ensuring balance within the data sets used for training and testing ML models [5]. This strategy aims to address representation bias due to the imbalance in the representation of older adults and subsequent evaluation bias against different age groups in the data set, which can lead to biased predictions and unfair outcomes [5]. By ensuring a more balanced distribution of samples across age groups, the model can learn from a more diverse and representative set of examples, reducing the potential for age-related bias. A total of 4 papers used data set balancing techniques to achieve an equitable distribution of data across different classes or categories within the data sets, such as altering their data set or creating a new data set (as was the case for Jung et al [35]) that would balance their previously unrepresentative data.

For their study to demonstrate the effect of bias on older adults with dementia on facial expression analysis models, Taati et al [39] developed a data set of test participants comprised entirely of older adults for their study. This data set was comprised of images of 86 older adults (aged 65 years or older), of whom 42 were affected by dementia and 44 were cognitively healthy, creating a balanced data set with the target population [39]. Frontal and profile photos of each participant’s face were taken at baseline, and a physiotherapist guided each participant through a series of exercises to identify painful positions. Images were then annotated manually according to the facial action coding system (FACS) and the “Pain Assessment Checklist for Seniors with Limited Ability to Communicate-II” (PACSLAC-II) pain scales. This method helped avoid potential representation biases that could arise from an imbalanced representation of different cognitive states. Taati et al [39] also used a fine-tuning method, in which they pretrained their models with images of cognitively healthy older adults and then fine-tuned their models with images of cognitively impaired older adults. They found that when this strategy was tested on active appearance models (AAMs), the number of images in the fine-tuning data set with a normalized root-mean square error (NRMSE) of <5% improved from 87% to 91% accuracy. However, when fine-tuned with the same strategy, facial alignment networks (FANs) performance remained around 90% (for an NRMSE threshold of 5%). When the NRMSE threshold was lowered to 4%, the performance disparity became even more significant: AAMs and FANs both started between 65% to 70%, but the number of fine-tuning images with an NRMSE of <4% continually increased into the 75% to 80% range when the AAM was fine-tuned, while FANs did not see any increase (second and third figures in Taati et al [39]). The gaps in performance between AAMs and FANs indicate that the bias present in these results is at least partially algorithmic in nature, as both models were tested using the same strategy but only the AAMs showed any improvement [5].

In another paper, Zou et al [38] tested a model intended for cost-sensitive facial age estimation using the FG-Net and MORPH data sets, along with an image database of 14,238
images taken from Wikipedia Commons. They modified the FG-Net data set by adding images from their Wikipedia data set to balance it for age, particularly for the groups aged between 40 and 49 years and between 50 and 59 years. Doing this corrected the representation bias in the FG-Net database, 87% of which consists of participants aged 30 years or younger [38]. This mitigation strategy effectively resulted in a significant reduction in mean absolute error (MAE) for those age groups, along with a smaller reduction in MAE for the other age groups in their test. The balanced data set using the cost-sensitive function showed the lowest MAE in age predictions (MAE 8.25, SD 0.03) versus the cost-insensitive data set (MAE 9.31, SD 0.4) and an unmodified data set (MAE 8.6). This approach effectively reduced the MAE for the groups aged between 40 and 49 years and between 50 and 59 years, so it was similar to the younger groups aged between 20 and 29 years and between 30 and 39 years. However, the representation of the groups aged between 60 and 69 years and 70 and 79 years appear to have been largely unchanged (fifth figure in Zou et al [38]), with every other age category receiving a substantial number of additional images except for the 2 oldest, and as a result, the MAE for these groups seems to be substantially higher, even using the cost-sensitive model (eg, for the group aged between 60 and 69 years, the MAE was 17, and for the group aged between 70 and 79 years, the MAE was 30), while the younger age groups did not have an MAE above 10 [38].

Jung et al [35] worked with several databases on facial recognition, including the IMDB-Wiki data set, a data set composed of Twitter profile images, and the 100 Celebrities data set. Jung et al [35] created the 100 Celebrities data set after noticing that the other data sets were imbalanced for age and ethnicity. Celebrity images were selected due to the wide availability of high-quality images of celebrities from a wide range of angles and the simplicity of establishing the participant’s true age when the photo was taken (celebrity birthdates are easily determined). Creating this data set would help mitigate the representation and evaluation bias found in the IMDB-Wiki data set, although using those data sets without balancing them would still expose the outcomes to those biases [5]. Age detection by the Face+++, IBM Bluemix Visual Recognition, AWS Rekognition, and Microsoft Azure Face API detection systems found that on the balanced 100 Celebrities data set, age was underestimated by 15.2 years. The highest accuracy of all the models trained on the 100 Celebrities data set was the IBM model, at 53% (although the Microsoft model achieved 66% accuracy when trained on the Twitter-Age data set). However, it is worth noting that the 100 Celebrities data set is comprised of celebrities who fall under entertainment-industry beauty standards, which may explain the large variance. While the creation of a balanced data set is a step toward mitigating biases, the fact that the age detection models still exhibited significant inaccuracies in this study raises questions about the overall effectiveness of the strategy with this specific data set. By grouping all older adults into a single category (aged 55 years or older), the 100 Celebrities data set may also increase the likelihood of aggregation bias, as mentioned in the previous section.

Finally, Liang et al [36] attempted to balance their combined data set that originally contained magnetic resonance imaging (MRI) scans of the brain from 2026 participants. These were comprised of samples from the ABIDE, CoRR, DBLS, and NIKI Rockland data sets that had a higher number of participants aged 40 years or younger. To help balance representation bias within this sample, the authors resampled the full data set in 5-year age intervals to address overrepresentation and underrepresentation of age groups. They duplicated samples from overrepresented age groups to match the number of samples in underrepresented age groups (ie, participants aged 75 years or younger), resulting in a more balanced distribution of age groups within the new data set (n=782), reducing the bias toward certain age ranges, and ensuring a more representative representation of the population [26]. However, the researchers noted that a significant bias persisted even after training their model on their more balanced data set. Testing their model on the imbalanced data set returned r=.91 and an MAE of 6.77 years, while the balanced data sample returned r=.91 and an MAE of 8.02 years. The correlation between the brain-age gap and chronological age remained the same (–0.52) for both the imbalanced full data set and the balanced data sample. This indicates that, despite accounting for representation bias by balancing the data set, the strategy alone did not achieve the desired reduction in bias. Moreover, the MAE increased from 6.7 years to 8 years; although the correlation remained the same, the higher MAE suggests a decrease in the accuracy of the model’s predictions when using the balanced data set. These findings suggest that additional strategies or factors may need to be considered to further mitigate bias and improve the accuracy of the model. It is possible that the resampling strategy, while addressing representation bias to some extent, may not have fully addressed other sources of bias present in the data set. Therefore, this strategy had limited effectiveness in addressing bias while aiming to improve accuracy.

Data Augmentation

Data augmentation strategies can mitigate bias related to ageism in ML models by enhancing the diversity and representativeness of the data set. This enables the model to learn from a wide range of age groups, reducing potential bias toward specific age categories. In 3 papers that applied data augmentation techniques, the actual images in their data set were modified without adding images from an external source. For example, Georgopoulos et al [34] applied digital age progression methods to the images in the data sets selected for their study, generating realistic images of what their participants may look like as they age but modified images of the participant. Smith and Ricanek [37] used random cropping and Gaussian tinting, and Abderrahmane et al [32] used an unspecified technique.

Using the MORPH, Cross-Age Celebrities data set (CACD), and FG-Net data sets, Georgopoulos et al [34] tested the ability of a generative adversarial network (GAN) to synthesize aging patterns realistically. They divided the participants’ images from each data set into 4 groups: those aged 30 years or younger, between 31 and 40 years, between 41 and 50 years, and 51 years or older. For each data set, their model would then take an image from these data sets and either age or de-age new images for each of the other 3 categories to which the participant did not
belong, creating new images for each of the other 3 age groups and balancing the data set in the process. The approach was effective in creating a data set 4 times the size of the original data set, and their results were able to significantly improve both the data set diversity (measured using the Shannon D and Simpson H metrics) and overall balance (measured using the Shannon E and Simpson E metrics) of the data sets they studied, demonstrating superior performance over contemporary models (Conditional Adversarial Auto Encoders [CAAEs] and Identity-Preserved Conditional Generative Adversarial Networks [IPCGANs]). While CAAEs generated the most accurate images for the group aged between 31 and 40 years after being trained on the MORPH data set (MAE 1.18), and IPCGANs generated the most accurate images for the demographic aged between 31 and 40 years after being trained on the CACD data set (MAE 0.04), the novel method presented in the paper was most accurate for the groups aged between 41 and 50 years and 51 years or older, for both the MORPH and CACD (MAE of 1.21 and 1.69, and MAE of 1.33 and 1.04, respectively). The images generated by the novel method also produced the best scores on the Simpson and Shannon data diversity indices after augmenting the MORPH, CACD, and FG-Net databases. This approach effectively reduced the representation and evaluation bias against older adults in these data sets. In doing so, the researchers acknowledged the possibility of their model enabling researchers to overcome demographic bias in facial image data sets, the most popular of which heavily underrepresent older adults [34]. The method used could also have the impact of increasing aggregation bias, as all older adults were grouped into a single group (aged 51 years or older) [5].

Smith and Ricanek [37] studied age and gender prediction models using data sets taken from IMDB, Wikipedia, and the MORPH data sets, which underrepresent older adults and present representation and evaluation bias [37]. They sought to expand the robustness of their data sets by applying random data-augmentation policies, which are transformation techniques used to modify existing data. For instance, they used random cropping and Gaussian tinting techniques to increase the diversity of the images in the data set. After training and testing, they also composed a separate challenge data set, which applied the data-augmentation policies to images from categories that their model had difficulty identifying accurately, including female individuals, older adults, and individuals with darker skin. The MORPH data set lacks images of older adults, so adding additional images was effective in addressing representation and evaluation bias in the data set [5]. The data augmentation policies would be randomly applied to these images as they were loaded during training [37]. This method to augment the training data was effective, as the overall MAE fell from 4.62 to 4.21, with a final MAE of 4.13 for male individuals and 4.90 for female individuals. The overall gender prediction accuracy increased slightly, from 98.44% to 98.92%.

Abderrahmane et al [32] developed an algorithm for age prediction based on hand images. They acknowledged that their data set initially had a highly imbalanced age distribution, which could introduce bias in their model. To address this, they used data augmentation techniques to create a more balanced data set from the 11,000 Hands Database, which contained 11,000 images of hands from 190 participants, showing the dorsal and palmar aspects. The authors recognized that the data set was underrepresentative of older adults and appeared to use data augmentation to balance the data set. There were no specific details regarding the augmentation processes used, but the figures suggest that additional images were added to the data set to address the underrepresentation of certain age groups. The underrepresentation of older adults in the 11,000 Hands data set is noteworthy in light of the significant use of hand images to represent older adults (more details on this issue have been provided in the Discussion section).

Algorithmic Alteration

Papers were listed under algorithmic alteration if the researchers adjusted the calculations of their algorithm itself and applied statistical methods to their algorithms to reduce the bias in their outcomes [5]. We found that 2 papers had adjusted their algorithms to improve performance [33,36]. Liang et al [36] applied linear regression to correct for the bias produced by their model after balancing the data set, which proved unsuccessful. Noticing a bias in age prediction based on MRI scans, which resulted in less accurate predictions for older participants, they unsuccessfully attempted to balance their combined data set, as was previously discussed. After examining multiple possible additional sources of the bias, including noise within the data, heterogeneity of the data sets, and the use of specific ML models, they were able to correct the bias statistically with a linear regression, fitting a linear regression model for predicted age to the chronological age and sex that improved MAE. By ruling out the other sources of bias (such as representation and evaluation bias, which they corrected for by balancing their data set), they were able to determine that the algorithm itself was the source of the issue (algorithmic bias, per the framework by Mehrabi et al [5]). The study by Liang et al [36] is unique in that they attempted to correct for more than one type of bias: first working on representation bias, and then solving for algorithmic bias (biases introduced by the algorithm itself) after balancing their data set was unsuccessful, which was an overall effective approach.

In the second paper, Clapes et al [33] sought to correct the bias between estimated age and true age by dividing participants into smaller groups based on mutually exclusive image categories and recalculating the estimated age for each real age. This produced a fitted curve for the difference between estimated and real ages, which was then used to correct the bias between the estimated and real ages. Linear interpolation was used to correct bias for ages that had fewer examples. This effectively addressed the measurement bias in the study. Clapes et al [33] also added labels for expression, ethnicity, makeup, gender, and the age of the photograph itself to the APPA real database. The resultant model’s overall performance improved, reducing the MAE from 13.57 to 12.07. The reduced representation of older adults is mirrored by trends in the consistency between age predictions in Clapes et al [33], where the difference between real age and estimated age widens as age increases (panels A-F in the eighth figure in Clapes et al [33]), which they theorized was due to the decreasing representation of older age groups (ie, representation bias).
Discussion

Overview

Our review included 8 publications that used mitigation strategies to address age-related bias. To our knowledge, this is the first review to examine this topic. Our analysis revealed that age-related bias predominantly stemmed from the underrepresentation of older adults in the data sets used to build the models (representation bias). Notwithstanding the 100 Celebrities data set, the data sets in this study (for which data were available) contained only 0.05% to 7% of data representing older adults. Our first research question explored the variety of strategies used to address this bias and their effectiveness. Researchers used three approaches: (1) creating a smaller yet more balanced data sample from their existing data set (n=4) [35,36,38,39], (2) augmenting and supplementing the available data (n=3) [32,34,37], and (3) modifying the algorithm directly to account for bias specifically (n=2) [33,36]. There was heterogeneity in the outcome measures, so a meta-analysis was not possible. Our findings emphasize the multifaceted nature of bias in ML models and the strategies available to address it, as well as the critical imperative of identifying and mitigating age-related bias in ML models to ensure fairness and equity for older adults in society [32-39].

The effectiveness of mitigation strategies aimed at reducing age-related bias in ML models varied based on several factors, including the types of data used, the ML approaches used, and the specific purpose of the ML model. While the included papers covered a range of uses, including facial and age recognition and MRI brain-scan interpretation, it is crucial for researchers to recognize that the solutions that may apply to one type of model may not apply to others due to qualitative differences in the data each model depends on. While many researchers reported successful outcomes with their methods, some encountered challenges and limitations. Jung et al [35] conducted experiments with a more balanced data set but found that the accuracy of their model reached a peak of only 53%. This suggests that achieving complete mitigation of age-related bias may be difficult, even with an improved data set balance. Taitt et al [39] explored a mitigation strategy that showed mixed results and discovered that adding images of older adults affected by dementia to the training data for models trained on cognitively healthy older adults improved landmark detection significantly, but this approach improved the performance of AAMs significantly more than FANs. This highlights how specific types of algorithms can affect outcomes and how that must be considered when selecting which models to use when attempting to reduce bias. In a follow-up study, Asgarian et al [40] also identified performance differences in models intended to identify facial landmarks between healthy older adults and older adults with dementia [46]. In another example, Liang et al [36] encountered challenges in their attempts to balance data set representation and ultimately settled on a linear regression alteration for their algorithm. Making algorithmic alterations appears to have been most effective at mitigating measurement bias, but this is an extremely small sample (n=2) [33,36]. Overall, these findings highlight the complexity and context-dependence of mitigating age-related bias in ML models. While some strategies showed promising results, achieving complete elimination of bias remains challenging, and alternative approaches may need to be explored. Researchers must consider how both the data and the models they are using may affect outcomes.

Our second research question explored the types of bias being mitigated. Our previous work has demonstrated that age-related bias is present across the ML life cycle [9]. Bias affects ML models at multiple levels, including the data that trains the models (data to algorithm), the models themselves (algorithm to user), and the people who rely on the models (user to data) [5]. Our results found that researchers who attempted to correct for bias primarily focused on representation bias (n=7), with algorithmic bias (n=2) and measurement bias (n=1) being far less common targets of mitigation efforts. Representation and evaluation bias and measurement bias were most commonly mitigated by data set balancing [35,36,38,39], but this method did not consistently achieve the desired reduction in bias [35,36], while algorithmic bias [33,36] was resolved by algorithmic alterations, to the satisfaction of their respective authors.

Moreover, we noted that 2 papers may have aggregation bias [34,35], which occurs when a demographic is grouped in a way that does not account for its heterogeneity [5]. Many data sets in this review grouped all older adults into arbitrarily large categories (eg, those who were aged 55 years or older and 60 years or older). For example, Georgopoulos et al [34] aggregated the oldest participants into a single category of those aged 50 years or older [34]. While this worked to address one type of bias (representation bias), it reinforced another type of bias (aggregation bias), which impacts older adults significantly and essentially erases older adults as a distinct group, given that the legal cutoff to be considered a senior citizen is usually 65 years old [5]. While the 100 Celebrities data set by Jung et al [35] offered older adults equal representation compared with other age groups, it also grouped all older adults into a single category (ie, aged 55 years or older). This point of discussion holds significance but is often not fully acknowledged when tackling bias in ML. It highlights that bias can manifest in various interconnected ways, and while addressing the most prominent and evident forms of bias remains crucial, it may not completely eradicate bias from a model. Much of the discussion of bias in ML focuses on representation bias of specific demographics (it was the most common type of bias in this review), but there are other forms of bias that may be more easily neglected (such as aggregation bias, which was not addressed in the included studies). Sometimes, efforts to mitigate bias did not reach the data set’s oldest demographics. Panel 5A from the fifth figure in Zou et al [38] shows that the authors only balanced the FG-Net data set, and while they balanced the representation for all other age groups, they did not significantly improve the representation of the groups aged between 60 and 69 years and 70 and 79 years, which remained underrepresented. There are also ageist contradictions that emerge when representation is examined, with the strongest example in these papers being how even the images of the hands of older people were underrepresented in the 11,000 Hands data set, despite the prominence of wrinkled, disembodied hands as a dominant social representation of older people in media and societal
images [47]. Such imagery may have a dehumanizing effect by portraying older adults as a pair of hands rather than a face or a whole person, which could result in their apparent exclusion from a data set of images with similar subject matter being compiled for practical purposes [48].

An important implication is for developers to consider the practical significance of real-world impact of the mitigation strategy. For example, Clapes et al [33] reduced their MAE from 13.6 to 12.1, while Smith and Ricanek [37] reduced their overall MAE from 4.6 to 4.2, with a final MAE of 4.1 for male individuals and 4.9 for female individuals. While these are quantifiable improvements, it is also important to consider the extent of an improvement and whether it has practical meaning or relevance in decision-making processes.

To address age-related bias in ML, researchers and designers can take steps from research design, methodology, and technical perspectives to ensure that technology is accessible and inclusive. From a research design perspective, an alternative mitigation strategy for developers collecting their data sets is to provide a representative sample of participants. Jung et al [35] created an age prediction training data set using celebrity images from IMDB. The researchers intended to create balanced data sets for their target populations, which aimed to prevent representation and evaluation bias, reducing the potential for age-related biases in the analysis and evaluation of the data. Although the effectiveness of this specific approach is not explicitly stated, the use of balanced data sets likely played a role in mitigating potential biases stemming from an imbalanced representation of age groups. Consequently, their study results may be more reliable and less susceptible to age-related biases, contributing to a more unbiased understanding of age recognition algorithms (although they may still be affected by other varieties of bias, as was previously discussed). By emphasizing the importance of representative sampling and the creation of balanced data sets, researchers can minimize the risk of introducing age-related biases into artificial intelligence systems and foster more accurate and equitable results.

Lastly, promoting the consideration of age as a variable of interest could help draw the attention of researchers and developers to digital ageism and disparities within their data sets. Recognizing age as a crucial factor requires a multidisciplinary approach that involves educational institutions, industry leaders, and policy makers. Within educational curricula, emphasizing the significance of critical demographic parameters can sensitize future engineers and computer scientists to the diverse needs of populations, including older adults. Technical approaches that use robust statistical approaches can begin to minimize age-related bias in ML. Techniques such as stratification and oversampling can ensure adequate representation of older adults in training data sets, attenuating the risk of underperformance or misclassification for this demographic group. These techniques are particularly relevant where the intersections of social locations of age and known underrepresentations of disability, race, and sexual orientation in ML models are concerned for 3 reasons. First, many of these types of bias (racism, sexism, and ageism) have their roots in similar problems, such as underrepresentation [4,9]. For example, a 2022 study of 21 age-recognition systems found that artificial intelligence systems consistently identified age with less accuracy across all age, gender, and ethnic categories, which the authors speculated was due to older adults being underrepresented in training data [49]. Second, the methods used to balance age-related bias in these papers may also have applications for other types of bias. For example, Sixta et al [50] reported a combination of strategies to reduce bias, including balancing their data for underrepresented demographics using external data and then further augmenting the data with different lighting and image quality adjustments. Third, the discussion around these types of bias in ML is much broader than the focus on digital ageism, making them an effective tool for drawing attention to these issues and advancing work in this area [9]. In a comparison of commercial facial emotional recognition (FER) systems, Kim et al [51] found that performance did improve for all demographics from 2019 to 2020, but that FER systems still performed best on younger adults and most poorly on older adults. The “black box” nature of these commercial systems, in which the method’s underlying algorithm is inaccessible, makes it impossible for external reviewers to assess the source of these improvements from an algorithmic standpoint [51]. However, digital ageism is not limited to ML and is a broader issue that stems from societal biases in the design, development, and deployment of ML models [9,16]. The multivalent nature of age-related bias in ML requires solutions from multiple sectors of society, including the public, private developers, government, and academic research [24]. Future research should focus on developing a global consensus of priorities that can mobilize the multitude of players from these sectors to advance knowledge about age-related bias and best practices to address this bias.

Limitations

One limitation of this study is the exclusion of publications in languages other than English, potentially excluding the viewpoints and solutions to these challenges found within other cultures. Also, our review would not have captured all examples of implicit age-related bias, only in papers that mentioned the keyword “bias.” However, the studies included in this review can serve as exemplars of implicit and explicit bias. The implicit nature of ageism in the context of the technology sector with limited ethical oversight and regulations underscores the importance of further research and policy development. Our review also only examined papers that attempted to mitigate age-related bias, not other types of bias (such as gender- or ethnicity-related biases). One major concern is that these biases and omissions may continue to produce exclusion and push older people whose experiences are not read as “youthful” further outside the peripheries of shared social and cultural everyday spaces, including (but not limited to) those mediated through technological systems. This is perhaps especially the case where age intersects with locations already known to experience bias, such as gender, race, and ethnicity.

Conclusion

This study explored and synthesized mitigation strategies for age-related bias in ML. The results underscore the value of 3 primary strategies for bias mitigation: data set balancing, data augmentation, and statistical techniques. However, the efficacy...
of these tactics demonstrated variability contingent upon factors such as data type, ML methodologies, and the intended purpose of the ML model. Although some researchers reported successful outcomes by diversifying their data sets, achieving complete elimination of bias remains challenging, and alternative approaches should be explored. The practical significance of the intervention should also be considered, as improvements in bias reduction may not always have substantial real-world impact, and further bias reduction and mitigation may still be required. A greater understanding of how digital ageism and age-related bias are propagated in ML and reproduced is needed across multiple sectors, including researchers and policy makers. Future research and policy agendas should include developing collaborative, comprehensive transdisciplinary strategies to foster fairness and inclusivity in the digital landscape.

Data Availability

Data sharing is not applicable to this article as no data sets were generated or analyzed during this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Extracted information. The following data sets appeared in publications included in this review, but subject demographic data was not readily available; thus, they were not included in the tables: 11k Hands, ABIDE, CoRR, DLBS, NKI Rockland, Pilots Parliament Benchmark, and a data set comprised of Twitter Profile data. [XLSX File (Microsoft Excel File), 17 KB - aging_v7i1e53564_app1.xlsx]

Multimedia Appendix 2

PRISMA Checklist. [PDF File (Adobe PDF File), 273 KB - aging_v7i1e53564_app2.pdf]

References


Abbreviations

- **AAM**: active appearance model
- **CAAE**: Conditional Adversarial Auto Encoder
- **CACD**: Cross-Age Celebrities data set
- **FACS**: facial action coding system
- **FAN**: facial alignment network
- **FER**: facial emotional recognition
- **GAN**: generative adversarial network
- **IPCGAN**: Identity-Preserved Conditional Generative Adversarial Network
- **MAE**: mean absolute error
- **ML**: machine learning
- **MRI**: magnetic resonance imaging
- **NRMSE**: normalized root-mean square error
- **PACSLAC-II**: Pain Assessment Checklist for Seniors with Limited Ability to Communicate-II
- **PRISMA-ScR**: Preferred Reporting for Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews

https://aging.jmir.org/2024/1/e53564

JMIR Aging 2024 | vol. 7 | e53564 | p.804

(page number not for citation purposes)
Abstract

In this article, we will provide a rationale for a web-assisted acceptance and commitment therapy (ACT) approach to loneliness among older adults, drawing upon theories from the literature on adult development and aging, emotion regulation, and loneliness. The intervention program was developed using the principles of ACT, which is a cognitive behavioral approach and unified model of human behavior change and psychological growth. The ACT intervention focuses on developing nonjudgmental present-focused awareness of internal experiences (thoughts, emotions, and memories) through strategies such as acceptance and mindfulness rather than directly modifying or removing them per se. The ACT intervention appears well-suited to assist older adults in coping with the challenges of aging, as the focus is on an individual’s willingness to sit with internal experiences out of one’s control (ie, acceptance), stepping back from negative or critical thoughts and developing greater kindness toward oneself (ie, defusion), discerning what is most important to one’s true self (ie, values), and building larger patterns of effective action based on such values (ie, committed action). The ACT intervention was developed as a resource for older adults who are socially isolated or having difficulty with social connectedness. Eight modules comprise the web-assisted ACT intervention program, which includes reading materials, video clips, and activities. Each module is followed by a summary, a homework assignment, a short quiz to assess learning, and a moderated discussion with a coach. The intervention program begins with reconnecting participants with their values. The goal of the ACT intervention program is to foster flexibility in a participant’s behavior so they can behave consistently with their chosen values, rather than becoming locked into a pattern of behavior that is driven by avoiding distress or discomfort. The ACT intervention approach is both novel and innovative, as it is based on ACT and leverages a behavioral health web platform that is flexible and inclusive in its design. The ACT intervention aims to help older adults become more socially connected, less lonely, and more satisfied with their relationships with other people. The emphasis that ACT places on values and living life in accordance with one’s values renders it an approach ideally suited to older adults. Finally, recommendations for future research regarding this approach to addressing loneliness among older adults is addressed.

KEYWORDS
acceptance and commitment therapy; psychotherapy; loneliness; technology; lonely; older adults; older adult; gerontology; geriatric; geriatrics; emotion regulation; mental health; elder; elderly; isolation; aging; mHealth; digital health; digital mental health; online health; online support; eHealth; internet; depression

Introduction

Loneliness, the subjective experience of feeling socially isolated, occurs at all stages of the life span [1]. Loneliness has been observed to impact older adults’ experience of depressive symptoms and lead to functional limitations [2,3]. In addition, loneliness has been found to impact older adults’ cognition, with loneliness positively associated with dementia [4-6]. Individuals who score high on loneliness visit physicians more frequently [7]. Loneliness has also been associated with cardiovascular risk and risk of mortality [3,8-10]. The crucial point from these studies is that across the life span, loneliness or lack of social relationships puts individuals at greater likelihood of mortality [11]. Older adults are an underserved population; the majority of older adults in need of mental health treatment do not receive care [12]. During the COVID-19 outbreak in early 2020, older adults were disproportionately affected with high rates of mortality and difficulty accessing telemedicine services [13]. COVID-19 has increased calls for the adoption of technology-based approaches to ensure older adults are able to access mental health services [14]. If older adults do receive treatment, it is typically pharmacological, despite the fact that many prefer counseling [15]. Reasons why older adults do not receive treatment include lack of knowledge about mental health
services, lack of perceived need for services, and stigma [16,17]. In addition, provider factors such as biases and misconceptions about aging also contribute to mental health services being underused by older adults [12,18].

To date, most research examining the effectiveness of psychosocial treatments with older adults has examined cognitive behavioral approaches. Studies have shown that these types of treatments are effective in treating depression in older adults in the community and primary care [19]. However, studies have produced mixed findings on their effectiveness with some subpopulations, such as depressed, medically ill, homebound older adults [20,21]. Cognitive behavioral therapy (CBT) approaches may also be suboptimal in the treatment of anxiety in later life [22,23]. Thus, there is a need to further develop and test psychosocial treatments to provide better care to older adults.

The purpose of this paper is to present an acceptance and commitment therapy (ACT)–based intervention program comprising a cognitive-behavioral web-assisted intervention to enhance social connectedness in older adults. We will review the literature on evidence-based programs for older adults and provide the theoretical basis for our ACT-based intervention program [24]. We will review the rationale for such an approach and provide a brief description of the ACT intervention program, as well as describe current evaluation efforts and future directions for the work.

**Ethical Considerations**

The study was approved by the New England Institutional Review Board (IRB tracking number: 120180244).

**Loneliness Interventions**

The deleterious impact of loneliness on physical and mental health points to the need for evidence-based intervention programs for older adults. Few evidence-based interventions have been developed for older adults, and creative approaches are needed. There have been various intervention approaches for addressing loneliness. Interventions approached at the individual level have included CBT, interpersonal therapies, and psychoanalytic therapies; interventions at the group level have focused on social skills training and groups centered around addressing shyness and depressive symptoms [25]. For older adults, interventions that may lessen the experience of loneliness have focused on social resources (eg, family and friends) and individual pursuits (eg, spending time on meaningful activities) [26].

Meta-analyses on the effectiveness of loneliness interventions have pointed to the effectiveness of psychological interventions [27,28]. A recent meta-analysis was conducted to specifically examine the effectiveness of psychological interventions in reducing loneliness [27]. The effectiveness of the various types of interventions did not significantly differ; however, reminiscence therapy had the largest effect size, followed by social identity interventions; CBT had the lowest effect size [27]. Similar results were observed in another meta-analysis where reminiscence therapy and social skills training had the largest effect size, followed by CBT [28]. However, there was only 1 study that used reminiscence therapy. In addition, the meta-analysis found higher effect sizes for intervention studies that (1) used the UCLA Loneliness Scale to assess loneliness, (2) used technology (telephone or computer) to deliver the intervention, and (3) addressed maladaptive social cognition [28]. This meta-analysis indicated that the use of technology (ie, telephone or computer) in delivering the intervention can be an effective way to address loneliness [28].

A major challenge in developing evidence-based programs for older adults has been access. Telehealth, the use of video conferencing to provide social support for older adults, is one approach that has been used during COVID-19 [29]. Interventions delivered through technology (ie, telephone or computer) have been found to be effective at reducing loneliness for older adults. One study examined the effectiveness of delivering a web-assisted intervention to older adults with chronic illness or handicap and found that computer and internet use significantly reduced loneliness among older adults [30]. In addition, lower mean levels of loneliness and depressive symptoms were found in a videoconferencing program for older adults in nursing homes after administering the intervention for 3 months [31]. Educational programs aimed at teaching older adults to use computers and the internet have also been found to significantly reduce loneliness [32]. In addition, a recent randomized controlled trial during COVID-19 examining the effectiveness of a web-assisted group intervention via Zoom found that it was effective in reducing loneliness and depressive symptoms among older adults [32]. A qualitative study examining the impact of a web-assisted intervention on loneliness among adults aged 18 to 64 years found that technology was viewed as a means of social connection. Web-assisted interventions have the ability to have a positive impact, as reviewed above. They can present information in a simple format (eg, video, graphics, audio), reach people in their homes and in rural areas, and reduce stigma related to seeking treatment. However, too much use of technology can lead to videoconferencing fatigue, and face-to-face communication is preferred [33].

**Cognitive Approach to Loneliness**

Theoretical perspectives on loneliness have encompassed a wide range of approaches. Weiss [34] presented an interactionist view of loneliness that focused on the types of relationship deficits in people’s social networks. He described 2 types of loneliness: emotional loneliness (absence of a close partner) and social loneliness (absence of friendships and community). Other scholars have advocated for a cognitive approach to loneliness, which emphasizes the discrepancy between individuals’ subjective perceptions of their social life and their desired quantity or quality of social relationships [35]. That is, loneliness is hypothesized to occur when individuals’ networks of social relationships do not meet their expectations [35]. Cognition can act as a mediator between perceived loneliness and the intensity of the experience. For example, loneliness is often increased or decreased based on one’s thoughts and beliefs about one’s social skills. If a person believes that they are awkward and drive other people away, they may view every failed social encounter as...
their fault and as something that cannot be improved in the future. Thus, the cognitive approach emphasizes causal attributions for social difficulties, as well as behavioral and personality traits.

Based on the cognitive theory of loneliness and past research examining the effectiveness of psychological interventions, a new approach to loneliness is warranted that uses technology to deliver the intervention components. The use of technology (ie, telephone or computer) has been observed to be an effective means of delivering an intervention [27,28]. Psychosocial interventions (ie, reminiscence therapy, social identity interventions, and CBT) have been observed to be effective in reducing the experience of loneliness [27,28]. A systematic review of communication technology interventions suggested a need for further studies in the fields of loneliness and web-based technologies to identify opportunities to reduce loneliness in older people [36]. We propose that ACT, an emerging evidence-based approach to the treatment of emotional distress, is a viable new intervention approach for addressing loneliness among older adults, delivered via computer or the internet [24].

**ACT-Based Intervention**

The intervention program was developed by drawing upon basic research on loneliness in older adults and using the principles of ACT. ACT is a cognitive-behavioral approach and unified model of human behavior change and psychological growth. ACT interventions focus on developing nonjudgmental present-focused awareness of internal experiences (thoughts, emotions, and memories); willingness to sit with internal experiences that are out of one’s control (ie, acceptance); stepping back from negative or critical thoughts and developing greater kindness toward oneself (ie, defusion); discerning what is most important to one’s true self (ie, values); and building larger patterns of effective action based on such values (ie, committed action). Compared to traditional CBT, ACT pays greater attention to the context and functions of private events and emphasizes helping individuals respond to them with greater flexibility through strategies such as acceptance and mindfulness rather than directly modifying or removing them per se.

There are several lines of evidence that support the use of ACT with older adults who are socially isolated or having difficulty with social connectedness. First, theories of adult development and successful aging suggest that an ACT approach to treatment could be useful for older adults [37]. It is well established that older adults experience a change in life dynamics due to shifts in gains and losses. Older adulthood has a higher proportion of losses and many of those are out of one’s control and unable to be changed [38]. Given that research has consistently shown that loneliness in older adults can be exacerbated by losses such as disability [39,40], decreased mobility [41-43], and widow- or widowerhood [43-45], it may prove fruitful to consider intervention approaches that assist older adults to respond flexibly to these relatively unmodifiable aspects of later life. The use of interventions was proposed to enhance psychological acceptance among older adults in response to findings of higher levels of well-being among older adults who showed greater psychological acceptance [46]. Also, in a study in which treatment response was defined as a reduction in the amount of interference with life due to pain, older people were more likely to respond to ACT compared to CBT [22].

Second, although CBT approaches have been used with some success, there is reason to doubt their effectiveness for all individuals. Traditional CBT strategies that involve challenging the validity of thoughts may not be beneficial to older adults because the thoughts and feelings that arise after losses, while unhelpful, may not be unrealistic. Moreover, because older adults often have beliefs about aging that have solidified over a long period of time, modifying them may not be an efficient use of time [47].

This is consistent with research on CBT for mental health problems, which has revealed that the cognitive restructuring components of CBT do not significantly improve therapeutic outcomes [48]. Such techniques may be invalidating, futile, or cause iatrogenic effects. For example, there are data to show that cognitions are not directly modifiable and that deliberate attempts to change or suppress thoughts can increase their occurrence and behavioral impact [49]. Furthermore, some evidence suggests that older adults who make active efforts to eliminate problems that cannot be solved are at a higher risk for depression and other negative outcomes, and that disengaging from commitments or goals that are unattainable—followed by choosing an attainable alternative—is associated with better emotional well-being [50,51]. An acceptance approach in which individuals learn to focus on their remaining resources may be more beneficial than an approach in which they are encouraged to modify their thinking about loss or disability.

Third, the goal of an ACT intervention is to live life in accordance with deeply held values. ACT may be particularly appropriate for older adults because individuals experiencing declines associated with aging may limit their goals to those that are most highly valued, work harder to strive toward achieving those goals, and use alternative strategies to compensate for formerly used strategies that may no longer be workable. This perspective is articulated in the Selective Optimization with Compensation Model of successful aging [52]. This is the essence of ACT and psychological flexibility. Such a model of treatment may resonate with older adults, which in turn may make it more likely they will actively engage in treatment. For example, there is some evidence to suggest that attrition rates may be lower among older adults treated with ACT when compared to those who received CBT [22].

In sum, ACT appears well-suited to assist older adults in coping with the challenges of aging, such as losses in functioning and changes in social connections. ACT has been shown to be effective for older adults in other areas, such as reducing depression and anxiety and improving symptoms of chronic pain [53-55]. ACT interventions have been adapted for web-assisted delivery and have been shown to be effective in teaching ACT-based skills to manage conditions such as overweight, fibromyalgia, and stress [56-58]. Recent research indicates even 4 web sessions can have a positive impact [59].
Development of a Web-Based ACT Intervention for Older Adults

Table 1 provides an overview of the ACT intervention program components. The ACT intervention program begins with reconnecting the client with their values, defined as intrinsic reinforcers, which provide a chosen direction for their behaviors and actions despite obstacles faced [60]. First, a list of relationships is provided (friendships, romantic or intimate relationships, family relationships, acquaintances or neighbors, and new people) and participants are asked to indicate how much they value each of those relationships. Importantly, indicating the level of importance means what the participants would like to see in their life, and does not necessarily mean that those relationships are currently matching up with that desired level. Next, a list of potential characteristics and qualities that people often value in their relationships is provided, and the client chooses the 5 that are most important to them. The list includes qualities such as being trustworthy, accepting, spontaneous, kind, open, sincere, forgiving, and loyal. The ultimate goal of the ACT intervention program is to foster flexibility in the participant’s behavior so they can behave consistently with their chosen values, rather than becoming locked into a pattern of behavior that is driven by avoiding distress or discomfort.

<table>
<thead>
<tr>
<th>Module</th>
<th>Purpose</th>
</tr>
</thead>
</table>
| (1) Introduction and values | • Education about loneliness and why it is important to address loneliness  
                              • Education about values  
                              • Identifying and connecting with values and what the older person wants relationships to look like, to enhance motivation |
| (2) Exploring coping | • Awareness of triggers and consequences  
                        • Identifying how one copes with distress  
                        • Evaluating the workability of one’s coping strategies |
| (3) Awareness of thoughts | • Learning how the mind works and how thoughts influence behavior  
                            • Learning how some people are often sensitive to certain social situations |
| (4) The impact of thoughts | • Identifying one’s own sticky thoughts and how they keep one stuck |
| (5) Changing the impact of one’s thoughts | • Stepping back from thoughts  
                                           • Labeling thoughts |
| (6) Changing the impact of one’s thoughts | • Acceptance  
                                           • Self-compassion |
| (7) Social stuck points | • Identifying one’s interpersonal patterns and identifying where one might need skills  
                           • Learning different skills for finding the relationships one wants |
| (8) Putting it all together | • Integrating education and skills from modules 1-7  
                              • Identifying barriers and setbacks and making a plan for skill use |
| Homework | • Self-assignments after each module on behavioral tasks to provide exposure to previously avoided situations, with the goal of increasing skill use and increasing adaptive thought patterns |

This intervention program is an interactive, dynamic web-based intervention for social isolation and loneliness that delivers ACT over 8 interactive modules. The web-assisted modules are largely self-paced. The ACT intervention includes reading materials, video clips, and activities; each module is followed by a summary and homework assignment, a short quiz to assess learning, and a moderated discussion with a coach. Participants also have access to supplementary materials and email reminders. Participants are instructed to work through the modules in sequential order, approximately one module per week. Each module begins with a brief review of the content from the previous module before introducing a new concept. The homework includes self-assigned behavioral activities to complete each week (e.g., “set a specific behavioral goal to engage in this week where you practice the skill learned in this week’s module”), and the participant is prompted to enter a description of the self-assignment after the module. Before starting the next module, participants are prompted to complete an “activity recap,” where they enter whether or not they completed the behavioral exercise, and to indicate their satisfaction with the exercise. Finally, they are asked to share any challenges that arose while engaging in the activity, and if this kept them from completing it. Participants are sent reminders if they have not logged on to the system for over a week.
The exercises in the first modules of the ACT intervention program attempt to increase clarity about what participants truly value in relationships and social connections, examine the ways in which current behaviors are either helping or hurting these values, and help participants begin to take concrete steps toward behaving consistently with their social values. This is important to set the stage for the rest of the program. All educational and skill-building activities in the ACT intervention program are linked to participants’ values, and the module content is aligned with living life in accordance with their own personally chosen values. For example, an increased willingness to experience one’s distressing thoughts and emotions (instead of trying to change them) is a skill that is explored as a way to facilitate progress in areas of one’s life that provide meaning and personal fulfillment. Participants’ self-identified values are also used to set values-consistent goals. For example, if they have an overall goal to have more friendships, their specific goal might be to participate in more social activities where they can meet new people. Most importantly, the homework assignments at the end of each module (ie, self-identified behavioral tasks) are based on what is personally meaningful for each participant. For example, one participant cited being open, friendly, and active as key values. Her homework assignment was to attend the community’s senior center game night and talk to a person she had not met previously. Fostering these patterns of committed action is highly important, as this is what helps the participant increase their ability to engage in behaviors consistent with their values and increase social connectedness.

Next, the participants reflect on their use of coping strategies (ie, avoidance and control-based strategies) and evaluate if these are helpful. The goal of this module is to help the participants recognize that avoidance often does not work, and in fact may be their primary problem. For example, isolating oneself at home to control the uncomfortable feelings that arise when attempting to meet new people is an avoidant and unhelpful coping strategy. To address maladaptive coping strategies, participants are guided through several activities to identify positive and active coping strategies. It is even better if their active coping strategies also involve building social connections.

The next several modules focus on thoughts, particularly how the mind works in social situations and how thoughts might influence participants’ behavior. These modules also emphasize how some people are vulnerable to experiencing certain thoughts more often, especially in social situations (eg, sensitivity to rejection). Participants identify “sticky thoughts,” which are thoughts that tend to stick around and control one’s behavior, even when they are unhelpful. For example, a sticky thought might be “Nobody wants to get to know me.” Participants are encouraged to increase awareness of how these thoughts get in the way of relationships with other people. Participants are then introduced to several strategies for changing the impact of their sticky thoughts, such as defusion (ie, stepping back from thoughts), acceptance, and self-compassion. One session is then devoted to recognizing social situations in which to use their newly learned skills (eg, when meeting someone new, making conversation with someone, or experiencing conflict with someone). In the final part of the program, participants reflect on their values and set further goals to live their lives according to these values, solving problems related to any perceived barriers.

Several adaptations were made to the program to ensure that it was appropriate for older adults. Changes in cognitive functioning with aging are not universal and older adults show extensive variability, but it is important to make programming flexible and inclusive for all learning abilities and types. Older adults may experience a decline in cognitive speed, working memory, selective attention, and fluid intelligence. Research indicates that for people with cognitive impairment, information should be presented slowly and with frequent repetitions and summaries.

In the ACT intervention program, information is presented in multiple ways, and participants are encouraged to notice what does and does not work for them. To accommodate for a decline in some aspects of executive functioning, such as memory, new information is presented in the context of the previous module’s material. Participants can revisit module content at any time, including readings, videos, and assignments. Phone prompts or alarms can remind participants in the program to carry out certain activities, such as behavioral homework tasks.

Conclusion

The desire for fulfilling social connections is a universal need and essential to well-being. Involvement in satisfying social relationships contributes to enhanced emotional and physical health throughout the life span. The fulfillment of this need looks different for older adults, as they navigate the developmental changes of later life and adjust to gains and losses that occur in their social networks. Furthermore, factors such as fear of intimacy, low self-esteem, and behavioral struggles such as difficulty with social skills may exacerbate feelings of loneliness and also make it more difficult to recover from loneliness.

This intervention aims to help participants become more socially connected, less lonely, and more satisfied with their relationships with other people. Researchers and clinicians collaborated on the ACT intervention program, leveraging a behavioral health platform to deliver the intervention. Most existing web-assisted interventions for loneliness that have been tested for efficacy are therapist-based and require extensive involvement by mental health professionals, and they are therefore neither anonymous nor self-directed [27]. Unlike therapist-assisted programs on the web, the ACT intervention program is an interactive, dynamic web-based intervention for social isolation and loneliness that delivers material over 8 modules and can be done largely at one’s own pace and anonymously. A trained wellness coach is available to chat (voice or web-assisted) after each module to assist the user with processing program content or answering questions.

The ACT intervention approach is both novel and innovative, as it is based on ACT and leverages a behavioral health web-assisted platform that is flexible and inclusive in its design. The session content of the ACT intervention program is reflective of current knowledge of loneliness presented in the literature, as well as the extant evidence on ACT for older adults.
The emphasis that ACT places on values and living life in accordance with one’s values renders it an approach ideally suited to older adults. Furthermore, some of the basic notions of ACT (eg, that pain and suffering are inevitable in human experience and that trying to avoid pain and suffering leads to problems) are highly compatible with this developmental stage of the life span and consistent with basic research on older adults and well-being.

Future studies should include randomized controlled trials of the ACT intervention program that include an attention control group, rater-blind assessments, and a systematic investigation of possible mechanisms of action (eg, acceptance). Measures of treatment adherence will need to be established, and interactions with coaches should be monitored and assessed for treatment fidelity and competency. Using mediator-moderator analyses, physical and cognitive functioning should be evaluated as moderators, and intrinsic motivation and homework adherence as mediators of the ACT intervention program’s effectiveness. Additional research may include determination of optimal “dosages” (frequency and duration) of the program for different subgroups of participants. Finally, an examination of treatment components via dismantling studies may reveal the program activities that are most potent in reducing loneliness. For example, the extent to which chat with coaches is essential to the program’s success will need to be studied.

Acknowledgments

UnitedHealthcare and American Association of Retired Persons (AARP) Services Inc supported this work. The conclusions were solely those of the authors. The funders played no role in the findings or conclusions drawn by the authors.

Conflicts of Interest

None declared.

References


https://aging.jmir.org/2024/1/e47943


**Abbreviations**

ACT: acceptance and commitment therapy

CBT: cognitive behavioral therapy
© Amie Zarling, Joseph Kim, Daniel Russell, Carolyn Cutrona. Originally published in JMIR Aging (https://aging.jmir.org), 22.4.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Corrigenda and Addenda

Correction: Detecting Anomalies in Daily Activity Routines of Older Persons in Single Resident Smart Homes: Proof-of-Concept Study

Zahraa Khais Shahid¹,², MSc; Saguna Saguna¹, BIS (Hons), MIT, PhD; Christer Åhlund¹, MSc, LicEng, PhD

¹Division of Computer Science, Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Skellefteå, Sweden
²Information Technology Department, Skellefteå Municipality, Skellefteå, Sweden

Corresponding Author:
Zahraa Khais Shahid, MSc
Division of Computer Science
Department of Computer Science, Electrical and Space Engineering
Luleå University of Technology
Forskargatan 1
Skellefteå, 931 77
Sweden
Phone: 46 704741624
Email: zahraa.shahid@ltu.se

Related Article:
Correction of: https://aging.jmir.org/2022/2/e28260 (JMIR Aging 2024;7:e58394) doi:10.2196/58394

In Detecting Anomalies in Daily Activity Routines of Older Persons in Single Resident Smart Homes: Proof-of-Concept Study (JMIR Aging 2022;5(2):e28260), the authors noticed one error.

In the originally published article, a duplication error occurred in Figure 5. The original Figure can be viewed in Multimedia Appendix 1.

Figure 5 has been corrected as follows:

Figure 5. Probability distribution with mean (SD) for 351 days in the kitchen between (a) 6:00 and 9:00 and (b) 11:00 and 13:00 in apartment 1.

(a) Between 06:00_09:00; mean 87.64; (SD 28.79) 351 days
(b) Between 11:00_13:00; mean 52.28; (SD 28.02) 355 days

The correction will appear in the online version of the paper on the JMIR Publications website on April 30, 2024, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.
Social Media Programs for Outreach and Recruitment Supporting Aging and Alzheimer Disease and Related Dementias Research: Longitudinal Descriptive Study

Anthony L Teano¹, MLA; Ashley Scott², MHA; Cassandra Gipson², MPH; Marilyn Albert², PhD; Corinne Pettigrew², PhD

¹Department of Geriatric Medicine and Gerontology, Johns Hopkins University School of Medicine, Baltimore, MD, United States
²Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Corresponding Author:
Corinne Pettigrew, PhD
Department of Neurology
Johns Hopkins University School of Medicine
550 N. Broadway St., Suite 415
Baltimore, MD, 21205
United States
Phone: 1 410 614 0363
Email: cpettigrew@jhmi.edu

Abstract

Background: Social media may be a useful method for research centers to deliver health messages, increase their visibility in the local community, and recruit study participants. Sharing examples of social media–based community outreach and educational programs, and evaluating their outcomes in this setting, is important for understanding whether these efforts have a measurable impact.

Objective: The aim of this study is to describe one center’s social media activities for community education on topics related to aging, memory loss, and Alzheimer disease and related dementias, and provide metrics related to recruitment into clinical research studies.

Methods: Several social media platforms were used, including Facebook, X (formerly Twitter), and YouTube. Objective assessments quantified monthly, based on each platform’s native dashboard, included the number of followers, number of posts, post reach and engagement, post impressions, and video views. The number of participants volunteering for research during this period was additionally tracked using a secure database. Educational material posted to social media most frequently included content developed by center staff, content from partner organizations, and news articles or resources featuring center researchers. Multiple educational programs were developed, including social media series, web-based talks, Twitter chats, and webinars. In more recent years, Facebook content was occasionally boosted to increase visibility in the local geographical region.

Results: Up to 4 years of page metrics demonstrated continuing growth in reaching social media audiences, as indicated by increases over time in the numbers of likes or followers on Facebook and X/Twitter and views of YouTube videos (growth trajectories). While Facebook reach and X/Twitter impression rates were reasonable, Facebook engagement rates were more modest. Months that included boosted Facebook posts resulted in a greater change in page followers and page likes, and higher reach and engagement rates (all \( P \leq 0.002 \)). Recruitment of participants into center-affiliated research studies increased during this time frame, particularly in response to boosted Facebook posts.

Conclusions: These data demonstrate that social media activities can provide meaningful community educational opportunities focused on Alzheimer disease and related dementias and have a measurable impact on the recruitment of participants into research studies. Additionally, this study highlights the importance of tracking outreach program outcomes for evaluating return on investment.

(JMIR Aging 2024;7:e51520) doi:10.2196/51520

KEYWORDS

education; social media; outreach; recruitment; Alzheimer’s disease; Alzheimer disease
Introduction

With the aging of the population, the number of individuals living with dementia globally is expected to triple by 2050 [1]. Disseminating reliable information on topics related to healthy aging and Alzheimer disease and related dementias (ADRD) is particularly important given modifiable lifestyle factors may play a role in dementia risk reduction [2], receiving an early diagnosis requires identifying the signs and symptoms of memory loss, and persons and families affected by dementia may benefit from information about ADRD. Connecting the public with education and resources related to these topics is aligned with national and international initiatives that emphasize the importance of healthy aging, dementia awareness, risk reduction, diagnosis and services, and caregiver support [3,4]. Social media content aimed at older adults may provide a mechanism for sharing resources, research findings, and available services related to healthy aging and ADRD. Social media may also generate increased interest in research participation, in addition to the in-person approaches that have been used for many years.

Research centers specializing in aging and ADRD use several approaches for community outreach, education, and volunteer recruitment for research studies. This includes a variety of in-person events, such as providing educational presentations in the community, hosting resource tables at community events, and organizing conferences in collaboration with community organizations, among others. Centers also use a range of web-based activities including newsletters, websites, and social media platforms such as Facebook, X (formerly Twitter), YouTube, and Instagram. While social media may be a potentially important method of delivering health messaging and recruitment [5], the types of activities used, and the outcomes of social media outreach efforts, are not well understood.

Social media has become a primary source of information for the general public hoping to “seek and share health-related information” [6]. Recent estimates indicate that 73% of middle-aged and 45% of older adults in the United States use social media, with Facebook and YouTube being two of the most common among these age groups [7,8]. This suggests that key age demographics for messaging may be reached through social media, providing a platform for research centers to become trusted web-based sources of information on topics related to healthy aging and memory loss, ADRD, and caregiving in their local communities.

Studies incorporating social media activities into community outreach and recruitment in the context of aging and memory loss have covered a broad range of topics. For example, studies have examined the role of social conversations in providing advice related to cognitive decline [9]; described social media activities in targeted dementia awareness campaigns [10-12]; and evaluated the use of social media advertisements to drive traffic to educational resources [13] or to recruit into aging or ADRD research studies [14-16]. Moreover, prior research has demonstrated the feasibility of using a digital community-engaged research approach (which included a social media component) for reaching individuals from racially diverse backgrounds [17]. More broadly, studies analyzing dementia content on social media have also suggested that these platforms have the potential to deliver health information on this topic, raising awareness and facilitating communication with the public [18-20].

Prior studies in this area have not, to our knowledge, described aging and memory loss educational programs across multiple social media platforms and reported on their impact on recruitment. To address this gap, this study describes the social media approaches, aimed at community outreach and education related to aging and ADRD, implemented by one research center. Importantly, the objective was not to compare social media to other approaches but to describe the social media activities implemented to date and evaluate the outcomes of these activities using up to 4 years of metrics, including success in reaching members of the community and impact on recruitment.

Methods

Overview

The Johns Hopkins Alzheimer’s Disease Research Center (JHADRC) has used 3 social media platforms over the past 4 years: Facebook, Twitter (rebranded as X in 2023), and YouTube. Each platform was launched with a different set of goals.

Facebook

The primary goal of the Johns Hopkins Memory & Aging Facebook page is to provide education and resources, as well as information about community events and research opportunities, related to healthy aging, brain health, memory loss, and dementia. The content primarily targets members of the community and community partners. Content is posted to this page 3-5 times per week. This page was launched in January 2019.

Twitter

The goals of the Johns Hopkins Memory & Aging Twitter account are 3-fold. The first goal is consistent with those of the Facebook page: to provide education, resources, and information about community events and research opportunities related to healthy aging, brain health, memory loss, and dementia. This includes promoting the visibility of Johns Hopkins (JH) aging and ADRD researchers, clinicians, and staff. The second goal is to serve as an information hub for local aging and ADRD researchers by sharing academically focused content such as recent research publications, information about social determinants of health for underserved populations (eg, Black Americans, Hispanics, and the LGBTQ+ community), funding and training opportunities, professional development opportunities, and upcoming conferences. Where possible, content is shared (ie, retweeted) through related JH Twitter accounts, such as the Center on Aging and Health and Geriatric Medicine and Gerontology, to expand views of the content. The third goal is to strategically network with community partners engaged in activities related to aging and ADRD. This includes making a concerted effort to share the events, activities, and
accomplishments of these groups. Given these 3 broad goals, the target audience includes community audiences and community partners, as well as researchers and health professionals. A communications specialist spends approximately 1-3 hours per weekday identifying, sharing, and retweeting content of relevance to the abovementioned audiences, and approximately 6-8 hours monthly creating original content. This account was launched in July 2020.

Readers should note that the Facebook and Twitter pages are branded with the name Johns Hopkins Memory & Aging. This name was selected to be intentionally broad in order to demonstrate relevancy to topics related to aging, memory loss, and ADRD, and to resonate with the general public, particularly the target audience of middle-aged and older adults.

**YouTube**

The primary goal of the Johns Hopkins ADRC YouTube page is to serve as a repository for outreach content, including center-produced resources, recordings of web-based community outreach events, and activities developed in collaboration with the JHADRC’s community advisory board, known as the Memory and Aging Community Advisory Board (MACAB). The page primarily targets members of the community, with content uploaded as needed. This page was launched in October 2020.

All social media pages are overseen by members of the JHADRC’s Outreach, Recruitment, and Engagement (ORE) Core. The ORE Core worked with the Johns Hopkins Medicine Strategic Marketing and Outreach team to set up the page and get approval for the use of the institutional name. The web-based recruitment materials of the JHADRC were approved by the Johns Hopkins Medicine institutional review board (IRB).

**Description of Content**

The majority of content shared through the Facebook and Twitter pages highlights topics broadly related to ADRD, consistent with the goals of these pages. This includes educational content related to healthy aging, brain health, and risk factors for memory loss; information about Alzheimer disease and other types of dementia; resources for caregivers; information about clinical research and research findings; social determinants of health; and local events of relevance to these topics (both our own and those of our community partners). During the first year of the COVID-19 pandemic, many posts included COVID-19–related health and support resources.

The shared information includes content developed by JH staff (described below) and content from external sources such as news articles, website pages, infographics, and blog posts from trusted outlets. For example, we frequently share content from the National Institute on Aging, the Alzheimer’s Association, our local Alzheimer’s Association Chapter, the American Association of Retired Persons (AARP), a website known as MindCrowd, and major news outlets. We also make a concerted effort to share news articles and web-based resources that quote or feature JH staff, researchers, or clinicians, as well as materials that highlight recent research findings from studies that include JH researchers. As noted above, our Twitter page also shares content of relevance to academic audiences and health professionals, such as funding and conference announcements and professional development opportunities.

**Internally Developed Social Media Content and Programs**

**Social Media Series**

To date, we have developed 3 educational series consisting of a collection of posts on specific topics, including general information about AD, brain health, and research awareness (see Table 1 for details). Posts within each series are branded with a hashtag, allowing the topic to be indexed and searchable on social media platforms. Although developing a series of posts requires time up-front, the approach provides a library of original content that can be released over several months or more, and are available to be reposted in the future. This also allows for content to undergo IRB review prior to use if needed (eg, Research Awareness Series). Having prepared content on-hand has been particularly helpful given the multiple responsibilities that compete for staff effort.
Web-Based Talks About Memory Loss and Brain Health on Social Media

We have developed brief web-based presentations designed to reach new audiences in the local community. Web-based presentations on the topics of memory loss (“What you should know about memory loss”) or brain health (“A brain healthy lifestyle matters for healthy aging”) are given by ORE Core Community Outreach Coordinators. Individuals register through Zoom and registrants must attend to receive the content (ie, presentations are not recorded). One week prior to the presentation date, modest funds (US $30-US $150) are used to boost the post on Facebook, targeting middle-aged and older adults in the Greater Baltimore area. These presentations, organized approximately every quarter, have been advertised 11 times to date, at various times of day (eg, noon “lunch & learn”; evening “dinner table discussion”). Typically, more individuals register than actually attend. On average, 19 individuals, representing approximately 25% of those registered, attend each presentation.

Educational Videos Posted to YouTube

A web-based talk series, titled Memory Matters, was developed to provide brief research updates on topics related to aging, memory loss, and ADRD. These approximately 15-minute presentations are designed to share recent research findings that may be of particular interest to lay audiences, such as sleep, hearing loss, and physical activity (see the “Growth Trajectories and Page Visibility” section for details). To ensure that the presentations are targeted to a lay audience, presenters are asked to include an overview slide, make their slide content as simple and clear as possible, include a summary slide that recaps main points, conclude with actionable takeaway messages, and share publicly available resources. They also receive a document entitled “Tips and Examples for Effective Science Communication to a Lay Audience.” Presenters are given a template for the first and last slide to provide uniformity to the talks within the series. To date, these talks have been given by junior faculty who are given feedback in advance of the video recording by center faculty (CP and MA). Thus, this talk series also provides an opportunity for science communication mentorship to junior investigators, including guidance on effectively communicating complex scientific topics with lay audiences. Talks within this series are branded with the hashtag #HMemoryMatters. Once recorded and edited, they are posted to YouTube and the JHADRC website and shared through Facebook and Twitter.

A Community Views (#CommunityViews) web-based interview series features short (6-7 minutes) one-on-one interviews with members of the community. The goal of these videos is to provide members of the lay community an opportunity to

Table 1. Overview of social media series about aging and ADRD.

<table>
<thead>
<tr>
<th>Series name (social media hashtag)</th>
<th>Goal of content and distribution timeline</th>
<th>Description of content</th>
<th>Posts, n</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alzheimer’s awareness series (#JHAzAwareness)</td>
<td>Increase awareness of ADb Increase awareness of memory loss in aging &amp; Alzheimer disease</td>
<td>Two categories of content: (1) education about AD (eg, definitions, prevalence, signs and symptoms, stages, effect on the brain, and importance of research) and (2) risk factors and symptom management (eg, risk factors, risk reduction, genetics, diagnosis, treatments, and caregiving)</td>
<td>31</td>
</tr>
<tr>
<td>Brain health series (#JHBrainHealthMatters)</td>
<td>Share expert recommendations and practical tips from the reports developed by the AARP’s Global Council on Brain Health</td>
<td>Content and infographics from Global Council on Brain Health reports: (1) The Brain-Sleep Connection; (2) The Brain-Body Connection; (3) Engage Your Brain; (4) Music on Our Minds; (5) The Brain-Heart Connection; (6) Brain Food; (7) The Real Deal on Brain Health Supplements; (8) Preserving Your Brain Health During Illness or Surgery; (9) Brain Health and Mental Well-Being; and (10) The Brain and Social Connectedness</td>
<td>89 (7-10 posts per topic)</td>
</tr>
<tr>
<td>Research awareness series (#JHMemoryResearch)</td>
<td>Increase awareness of the importance of memory loss research and ongoing research activities to facilitate recruiting participants into JHADRC-affiliated research studies</td>
<td>Topics covered: (1) goals of research and different types of research studies (observational, interventions, or clinical trials); (2) examples of research procedures and what’s learned (eg, brain imaging and fluid biomarkers); (3) importance of diversity in research; (4) examples of JHADRC-affiliated research studies; (5) benefits of research participation; (6) participant, staff, and researcher spotlights; (7) study recruitment flyers</td>
<td>86</td>
</tr>
</tbody>
</table>

aADRD: Alzheimer disease and related dementias.
bAD: Alzheimer disease.
cAARP: American Association of Retired Persons.
dJHADRC: Johns Hopkins Alzheimer’s Disease Research Center.
advocate for topics of personal importance to them. Members of the JHADRC ORE Core work with the interviewee to collaboratively develop a set of interview questions, as well as a short set of slides to support the interview content and provide uniformity to the talks within the series. Once recorded and edited, the interviews are posted on YouTube and shared through Facebook and Twitter. To date, the interviews have been with members of the MACAB. Two interviews have been recorded: the first on the importance of educating the youth about dementia and the second on empowering older adults and caregivers to maintain their brain through physical activity and self-care. These programs are ongoing.

Twitter Chats and Webinars
A series of Twitter chats and webinars have been developed to promote the goals of a coalition of stakeholders in the brain health and dementia community. This collaboration consists of 2 research centers internal to Johns Hopkins University (the JHADRC and the Johns Hopkins Alzheimer’s Disease Resource Center for Minority Aging Research), as well as external partners including the Global Council on Brain Health convened by the American Association of Retired Persons, the Greater Maryland Chapter of the Alzheimer’s Association, and 2 strategically targeted audience stakeholder organizations, both local alumnus chapters of the Delta Sigma Theta Sorority, Inc. Branded as #BrainMatters, the goal has been to develop dialog, engagement, and educational programs for sharing evidence-based information about brain health, health disparities, memory loss, and ADRD at the grassroots level. This approach aligns with the recognized importance of developing equitable community partnerships for reaching individuals from diverse backgrounds [5].

Prior to each scheduled event, the program organizers communicate via email and meet virtually (typically 3-5 times) to make collaborative decisions about program timing, format, content, and invited participants. For Twitter chats, this involves identifying the specific questions to be sequentially asked during the scheduled web-based conversation; questions are answered by both invited guests and the Twitter users at large. For webinars, this involves deciding on the topic and event flow (eg, speaker presentations vs moderated conversations). To date, Twitter chat topics have included (1) brain health, (2) achieving brain health equity, and (3) the state of research on memory loss and dementia, and webinar topics have included (4) dementia caregiving and caregivers and (5) the relationship between community, social connection, and cognitive decline. Event promotion occurs predominantly on Twitter (which allows us to tag guest speakers) and Facebook, and is shared by the coalition’s stakeholders. The product of the Twitter chats remains on Twitter indefinitely as a resource to the public, and webinars have been recorded and uploaded to YouTube, thus, generating a marketable product that can also be shared after the event. In 2023, the #BrainMatters leadership made the decision to pivot toward webinars over Twitter chats, as it was determined that this format would better reach the target audiences.

Overarching Strategies

Research Recruitment
When the Facebook and Twitter accounts were first launched, we strategically decided not to emphasize clinical research recruitment (eg, study flyers, information about ongoing studies). Our goal was to first build a modest base of followers and develop a presence as a source of information before discussing research participation. Information about ongoing clinical research studies was not consistently posted until approximately 3 years after the Facebook page was launched. To date, this has been accomplished primarily through our Research Awareness Series (described above and in Table 1).

Diversity and Inclusion
We take care to ensure the visuals we post are representative of individuals from diverse communities. This includes evaluating the pictures posted in association with existing website links, as well as making a concerted effort to ensure that stock photos reflect individuals from diverse racial and ethnic backgrounds.

Boosting Facebook Posts
We have boosted occasional Facebook posts to increase the reach of our content and enhance the local visibility of our efforts, our Facebook page, and our center more broadly. This involves paying Facebook to more prominently display specific posts in users’ feeds. To help reach new local audiences who may be interested in our content, boosted posts targeted middle-aged and older adults (eg, 55 years and older) in the Greater Baltimore area (ie, approximately 25- to 50-mile radius of Baltimore, MD) using the Age and Locations fields of Meta Business Suite’s Boost Post settings. This strategy relied on Meta’s “Audience” settings, which allow users to define who will see their advertised posts by targeting audiences whose profiles match specific characteristics, such as demographics. Each time content is boosted, we make a concerted effort to invite individuals who like the posts to like or follow our Facebook page, to help expand our followership. Boosted content has most frequently included talks on social media, as well as posts from our educational series (eg, the Research Awareness Series) and information about center-organized events (eg, the MACAB’s Annual Holistic Health Seminar on Memory Loss). To date, our expenditures have been modest, ranging from US $30 to US $200 per boosted post. In years 2 and 3, we boosted 1 post approximately every other month; from year 4 onwards, we boosted 1 post almost every month.

Measures and Outcomes
In order to track our social media activities, at the start of each new month, members of the ORE Core record metrics from the prior month using data from each platform’s native analytics dashboards.

For Facebook, the following metrics were tracked through the Meta Business Suite’s Professional Dashboard: number of posts made during the month; the average reach of that month’s posts, reflecting the number of people who saw that month’s post at least once (calculated as the total [sum] reach of all posts within the month divided by the total number of posts); and the average
engagement with that month’s posts, reflecting the number of times people engaged with that month’s posts through reactions, comments, shares, and clicks (calculated as the total [sum] engagement with that month’s posts divided by the total number of posts). Note that all posts within a given month were included in these monthly analytics; therefore, the data reflect both organic and paid reach and engagement. The number of page likes and page followers was additionally recorded. Facebook metrics were missing for 2 months in year 2.

For Twitter, the following metrics were tracked through Twitter’s monthly Analytics report: number of Tweets; Tweet impressions, reflecting the number of times that month’s Tweets were displayed to users; profile visits; mentions; and new followers. The number of account followers and the number of accounts that we follow were additionally recorded. Two Twitter metrics (the number of account followers and the number of accounts we follow) were missing for 5 months in year 1, 2 months in year 2, and 2 months in year 3.

In this report, we summarize metrics for both Facebook and Twitter. This includes monthly data on the number of likes and followers to assess page growth trajectories (presented in 6-month intervals), which has previously been described as an indicator of success [5]. Monthly metrics were additionally used to calculate average reach rates and average engagement rates for Facebook, and average impression rates for Twitter. These were calculated by dividing each month’s average reach, average engagement, or average impressions by that month’s total number of page followers, then multiplying the quotient by 100. Annual reach, engagement, and impression rates were calculated for all years (excluding year 1) by averaging over all available months within a year, for year 2 onwards. Year 1 metrics were excluded because they appear inflated due to high engagement from a limited number of followers, as the pages were being established and accruing audiences. Note that the analytics for posts made at the end of the month may be slightly underestimated, given posts may have received continued views after the data were recorded. For videos posted to YouTube, we report the number of views over time.

The number of individuals who expressed interest in participating in research after engaging with our social media activities was tracked through a secure, web-based REDCap database hosted at Johns Hopkins [21,22]. This database is designed to track the outreach activities that result in recruitment, including how an individual learned about our studies, whether they were referred to a center-affiliated research study, and the outcomes of those referrals. Database entries that provide results related to social media come from 2 sources of information. The first reflects individuals who complete an IRB-approved “Participate in Research” webform on the JHADRC website and indicate on that form that they learned about us through social media. The second pertains to entries of individuals who contact us (eg, by phone) to express interest in research after viewing web-based programming, indicating that they learned about our programs through social media while completing their phone screen; these individuals were entered into the database by members of the ORE Core.

Statistical Analysis
The data reported below reflect outcomes and metrics between page launch (as noted by the dates above) and the time of manuscript writing (April 2023). Time-based outcomes (such as growth trajectories and video views) are reported as months since launch, to describe, for example, how long the content or programs have been available. Where relevant, continuous variables are summarized as mean (SD) and categorical data are summarized as n (%). Average annual reach, engagement, and impression rates are calculated as percentages (as described above), with differences between months with versus without boosted posts assessed by Mann-Whitney U tests. Analyses were run in SPSS (IBM Corp; version 29.0).

Social media data pertaining to this report are available by qualified investigators upon request to the corresponding author.

Ethical Considerations
The data for this report come from 2 recruitment procedures implemented by the JHADRC, both of which have been approved by the Johns Hopkins Medicine IRB (protocol NA_00045104). First, the JHADRC website includes a section labeled “Participate in Research.” This section includes a web form that asks individuals interested in learning more about research to complete an electronic form. The first item on the webform asks that the individual respond “yes” or “no” to the following statement: “I understand that by submitting this form I am providing consent to be contacted about research studies related to aging and memory loss by staff affiliated with the Johns Hopkins Alzheimer’s Disease Research Center, and that this information will be stored in a secure database.” Second, members of the JHADRC also administer an IRB-approved phone screen, which includes questions similar to those in the web form. The first item on the phone screen includes the following statement “Before we begin, I should point out that the possible risk to your answering questions is that you will be revealing confidential information to us. We promise to keep all such information strictly confidential. You do not have to answer any questions that might make you uncomfortable. Your participation is voluntary. If you do not agree, this will not affect your care, if you are a patient at Johns Hopkins, in any way. If you are not interested, please let me know.”

The anonymized, group-level demographic data included in this study are based only on outcomes from individuals who responded “yes” to the statement at the beginning of the web form or completed the phone screen and indicated willingness to be contacted about research. Participants were not compensated for completing these optional forms and outcome data are reported in aggregate to protect participant privacy and confidentiality.

Results

Outcomes and Indicators of Success

Growth Trajectories and Page Visibility

Growth trajectories for our Facebook and Twitter pages are shown in Figure 1. Both accounts have demonstrated continued increases in the number of new followers since the account launch, suggesting that it takes years to build a modest audience. The total number of Facebook followers exceeds 500; this
includes a clear uptick in the trajectory of Facebook page followers around 24 months when we started boosting occasional Facebook posts. Notably, months that did include a boosted Facebook post resulted in a greater change in page followers and page likes (mean change in page followers 13.7, SD 9.2; mean change in page likes 19.0, SD 12.3), compared with months that did not include boosted posts (mean change in page followers 6.2, SD 4.5, $U=294.5, P<.001$; mean change in page likes 8.0, SD 5.0, $U=285.0, P=.002$), suggesting our efforts to invite individuals to like or follow the page were successful. The total number of Twitter followers exceeds 2000. Additionally, the talks within the Memory Matters web-based talk series have received over 3000 views, whereas views of the Community Views interviews have been less (Table 2); both of the #BrainMatters webinars hosted to date attracted >100 registrants with >50 individuals attending.

Average annual reach and engagement rates (Facebook) and average annual impression rates (Twitter) for year 2 onward are shown in Table 3. As anticipated, for Facebook, average reach and engagement rates were significantly higher for months that included boosted posts (mean reach rate 34.7%, SD 26%, $U=332.0, P<.001$; mean engagement rate 5.1%, SD 4.3%, $U=297.0, P<.001$) compared with months that did not include boosted posts (mean reach rate 34.7%, SD 26%, $U=332.0, P<.001$; mean engagement rate 5.1%, SD 4.3%, $U=297.0, P<.001$).

**Figure 1.** Growth trajectories in social media followers for Facebook (left; launched January 2019) and Twitter (right; launched July 2020), shown in 6-month intervals.

### Table 2. Overview of educational videos posted to YouTube, including Memory Matters web-based talk series (#JHMemoryMatters) and Community Views (#CommunityViews) interview series.

<table>
<thead>
<tr>
<th>Talk title</th>
<th>Months since launch, n</th>
<th>Views to date, n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Memory Matters web-based talk series</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dementia and cognitive decline: a brief overview</td>
<td>42</td>
<td>658</td>
</tr>
<tr>
<td>Vascular risk factors for cognitive decline and dementia</td>
<td>41</td>
<td>430</td>
</tr>
<tr>
<td>Changes in sleep and circadian rhythms in aging and memory loss</td>
<td>22</td>
<td>414</td>
</tr>
<tr>
<td>Hearing loss and dementia: what’s the connection?</td>
<td>17</td>
<td>1629</td>
</tr>
<tr>
<td>The importance of brain donation for Alzheimer’s disease and related disorders</td>
<td>5</td>
<td>90</td>
</tr>
<tr>
<td>Physical activity and cognitive health</td>
<td>1</td>
<td>74</td>
</tr>
<tr>
<td><strong>Community Views interview series</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educating the youth about dementia</td>
<td>24</td>
<td>95</td>
</tr>
<tr>
<td>Empowering older adults and caregivers</td>
<td>3</td>
<td>35</td>
</tr>
</tbody>
</table>
Table 3. Visibility metrics over time for Facebook and Twitter. Values reflect mean (SD) and ranges.

| Yeara | Facebook | | | Twitter | | |
|---|---|---|---|---|---|
| | Average annual reach rate | Average annual engagement rate | Average annual impression rate | | |
| | Mean (SD), % | Range | Mean (SD), % | Range | Mean (SD), % | Range |
| All years (except year 1) | 80.1 (79.1) | 8.0-355.3 | 7.6 (5.8) | 1.2-29.2 | 1260.3 (1076) | 559.4-4152.9 |
| Year 2 | 50.2 (30) | 26.7-113.9 | 8.2 (4.7) | 4.1-16.5 | 1564.9 (1319.6) | 559.4-4152.9 |
| Year 3 | 52.5 (37.7) | 13.5-128.9 | 5.4 (2.7) | 2.1-11.2 | 803.4 (195.6) | 573.5-1141.6 |
| Year 4 | 138.4 (111.4) | 11.6-355.3 | 10 (8.3) | 1.4-29.2 | — | b |
| Year 5 | 62.8 (61) | 8.0-122.5 | 5.7 (5.3) | 1.2-12.0 | — | — |

aMetrics exclude each platform’s first year of data because these metrics appear inflated due to high engagement from a limited number of followers, as the pages were being established and accruing followers. For Facebook (launched in January 2019), year 2: January 2020-December 2020; year 3: January 2021-December 2021; year 4: January 2022-December 2022; year 5 (partial, n=4 months): January 2023-April 2023. For Twitter (launched in July 2020), year 2: August 2021-July 2021; year 3 (partial, n=8 months): August 2022-March 2023.
bNot available.

Research Recruitment Resulting From Social Media Activities

Social media activities have resulted in 89 individuals expressing interest in participating in center-affiliated clinical research studies over the past 2 years, the majority coinciding with boosted social media content (eg, posts from the #JHMemoryResearch series or web-based talks advertised on social media). This includes 76 individuals who indicated that they came to the JHADRC website after viewing social media content, when completing the “Participate in Research” website form (mean age 66.2, SD 10.4 years; 67/76, 88% female; 12/76, 16% self-reporting non-White race or Hispanic/Latino ethnicity). Of these, 46 (61%) have been referred to an ongoing research study, 22 (29%) were not referred, and 8 (11%) have phone screens pending (see Table 4 for additional details). In addition, 13 individuals expressed interest in participating in research after attending a web-based talk advertised on social media (mean age 64.3, SD 4.4 years; 13/13, 100% female; 3/13, 23% self-reporting non-White race). Of these, 12 (92%) have been referred to an ongoing research study; 1 (8%) was not referred (Table 4).

Table 4. Recruitment outcomes: breakdown of participant referrals for those expressing interest in participating in research after learning about the Johns Hopkins Alzheimer’s Disease Research Center through social media activities.

<table>
<thead>
<tr>
<th>Referred to a center-affiliated research study</th>
<th>Not referred to a center-affiliated research study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrolled in a center-affiliated research study</td>
<td>6 (8)</td>
</tr>
<tr>
<td>Enrollment pending (referral sent, outcome pending)</td>
<td>26 (34)</td>
</tr>
<tr>
<td>Not enrolled (eg, ineligible; study coordinator unable to contact)</td>
<td>9 (12)</td>
</tr>
<tr>
<td>Referred to another ADRC or online registries (eg, lives out of state; unable to travel)</td>
<td>5 (7)</td>
</tr>
<tr>
<td>Phone screen pending</td>
<td>8 (11)</td>
</tr>
<tr>
<td>Contacted but no longer interested</td>
<td>3 (4)</td>
</tr>
<tr>
<td>Unable to contact</td>
<td>19 (25)</td>
</tr>
</tbody>
</table>

Discussion

Principal Findings

This study describes one center’s social media strategy for community outreach and recruitment. Four years of page metrics and recruitment data suggest that the content and programs provide educational resources, increase the visibility of the center’s activities, and result in the recruitment of participants into center-affiliated research studies.

Given the increased use of social media among middle-aged and older adults, these platforms can serve as one method by which the JHADRC provides reliable education and resources to members of the local community, networks (on the web) with local community organizations, and shares information about the important work (research studies, research findings, events, etc) being done by center-affiliated faculty and staff. Based on these data, it appears that these efforts, including boosting occasional Facebook posts, have increased the visibility of the JHADRC in the local geographic area and have provided...
opportunities to reach individuals who we may not have encountered through in-person community outreach. Importantly, these activities can be achieved primarily through staff time, with minimal additional cost to center budgets. Using creative digital approaches to serve as a public source of reliable information on topics related to aging, memory loss, and ADRD provides an additional opportunity for centers to meet critical community engagement goals, including building trust and creating better communication with members of the community [23,24].

These findings also demonstrate that social media activities have been a source of participant recruitment. Potential participants (ie, middle-aged and older adults) learned about our research program through social media–advertised talks as well as social media post content, including our Research Awareness Series. This allowed us to increase the number of participant referrals to center-affiliated studies, supporting the promise of social media as a low-cost method of recruiting potential research participants [16,25,26]. Additional follow-up is needed to determine whether these participants have different demographics and retention characteristics than individuals recruited through in-person community outreach or other activities.

Our evidence suggests that innovative social media activities can also provide novel opportunities for the scientific community. For example, although our web-based Memory Matters talk series was primarily designed to share recent research findings with the lay community, it simultaneously provided professional development opportunities. Through this series, junior investigators receive mentorship on the principles and importance of science communication, specifically communicating science-related topics with the public [27,28]. Engaging staff and investigators in the development of social media content, such as writing accessible descriptions of their research programs and findings, can also provide science communication opportunities. This highlights an additional possible benefit of social media for centers or programs that may be considering whether to develop a social media presence or those exploring ways in which they might expand their programming.

Moreover, providing education and resources for adults in the local community, including individuals living with cognitive impairment and caregivers, is in alignment with international efforts promoting healthy aging, ADRD education and support, and dementia risk reduction [3,4]. While most of our educational content to date focused on these broad topics, future content could include materials that address ageism and ableism, and promote age-friendly [29] and dementia-friendly [30] activities and resources that ensure inclusive environments, empowerment, and sustained engagement with issues related to age and disease.

**Limitations**

This study has limitations. First, this is a retrospective account of one center’s social media activities and their outcomes, and the number of individuals following the social media pages is still modest. Additional work is needed to understand the long-term impact of these activities on community education, center visibility, and clinical research recruitment, and to evaluate the types of content that lead individuals to express interest in participating in center-affiliated research studies. Second, we do not have data on the demographics of our social media followers to evaluate whether our content reached the target audience or the extent to which we are reaching individuals living with cognitive impairment. However, the recruitment metrics indicate that those expressing interest in participating in research were in the target age range. Social media and internet use is reportedly lower among the oldest-old, as well as individuals with cognitive impairment and poorer subjective health [31-36]; these groups may, therefore, be less likely to engage with these platforms. Third, we did not compare the efficacy of the 3 social media platforms described, although we believe that each allows us to reach unique audiences in different ways. Fourth, the individuals who have expressed interest in research have been predominantly female; additional efforts are needed to understand how to similarly reach males, as well as a greater proportion of individuals from diverse racial and ethnic backgrounds. Fifth, although our reach and impression rates are promising, our Facebook engagement rates are relatively low. Because this metric is an indicator of how, and the degree to which, our audience interacts with our content, these data suggest that much of our social media outreach may be unidirectional (ie, low levels of likes, comments, and content sharing). Future efforts should strive to improve this. Additionally, these measures were collected as monthly averages, which limited our ability to assess the impact of individual posts or specific content types. Furthermore, engagement data were not collected for Twitter. Finally, we describe the efforts of one Alzheimer’s Disease Research Center; the extent to which similar approaches generalize to other types of centers or programs remains to be determined.

Although social media may be an effective means for raising awareness about dementia, dispelling stigma, and highlighting positive aspects of clinical research [37], it is only one method within a wide range of strategies needed for engaging communities, individuals with cognitive impairment, and families on topics related to aging and ADRD [5,24]. Social media may supplement in-person community outreach, engagement, and partnerships, the latter of which have been reported to be important for building trust, particularly in minoritized communities [38-40]. It will furthermore be important to continually monitor use trends among target demographics, and evaluate the addition of, or migration to, other emerging platforms. We, nonetheless, hope these activities—or a subset thereof—may serve as an exemplar for other centers or programs, or provide ideas for new initiatives that can be implemented and improved upon.

**Conclusions**

These data suggest that social media activities may have a measurable impact on the outreach, visibility, and recruitment activities of research centers, including National Institute on Aging–funded ADRCs. They also highlight the importance of tracking the success of outreach programs for evaluating outcomes. These data provide evidence of return on investment and support the continued use of social media for the above-mentioned purposes. Given the public’s use of social media as a source of health information [6], this may be an
important means by which centers can present themselves as a reliable educational resource, educate the community, and share research findings with community stakeholders. Incorporating additional activities designed to test the efficacy of different approaches for attracting research participants is an important future direction.

Acknowledgments
This work was supported by funding from the National Institutes on Aging (grant P30-AG066507).

Conflicts of Interest
None declared.

References
27. Feliú-Mójer MI. Effective communication, better science. Scientific American. 2015. URL: https://blogs.scientificamerican.com/guest-blog/effective-communication-better-science/ [accessed 2023-05-09]
Abbreviations

ADRD: Alzheimer disease and related dementias
IRB: institutional review board
JH: Johns Hopkins
JHADRC: Johns Hopkins Alzheimer’s Disease Research Center
MACAB: Memory and Aging Community Advisory Board
ORE: Outreach, Recruitment, and Engagement

©Anthony L Teano, Ashley Scott, Cassandra Gipson, Marilyn Albert, Corinne Pettigrew. Originally published in JMIR Aging (https://aging.jmir.org), 09.07.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Aging, is properly cited. The complete bibliographic information, a link to the original publication on https://aging.jmir.org, as well as this copyright and license information must be included.
Combating Barriers to the Development of a Patient-Oriented Frailty Website

Brian Greeley, PhD; Sally Seohyeon Chung, BSc; Lorraine Graves; Xiaowei Song, PhD
Fraser Health Authority, Surrey, BC, Canada

Corresponding Author:
Xiaowei Song, PhD

Abstract

This viewpoint article, which represents the opinions of the authors, discusses the barriers to developing a patient-oriented frailty website and potential solutions. A patient-oriented frailty website is a health resource where community-dwelling older adults can navigate to and answer a series of health-related questions to receive a frailty score and health summary. This information could then be shared with health care professionals to help with the understanding of health status prior to acute illness, as well as to screen and identify older adult individuals for frailty. Our viewpoints were drawn from 2 discussion sessions that included caregivers and care providers, as well as community-dwelling older adults. We found that barriers to a patient-oriented frailty website include, but are not limited to, its inherent restrictiveness to frail persons, concerns over data privacy, time commitment worries, and the need for health and lifestyle resources in addition to an assessment summary. For each barrier, we discuss potential solutions and caveats to those solutions, including assistance from caregivers, hosting the website on a trusted source, reducing the number of health questions that need to be answered, and providing resources tailored to each users’ responses, respectively. In addition to screening and identifying frail older adults, a patient-oriented frailty website will help promote healthy aging in nonfrail adults, encourage aging in place, support real-time monitoring, and enable personalized and preventative care.

(JMIR Aging 2024;7:e53098) doi:10.2196/53098

KEYWORDS
frailty; frailty website; patient-oriented assessment; community-dwelling older adults; internet security; privacy; barrier; barriers; development; implementation; patient-oriented; internet; virtual health resource; community dwelling; older adult; older adults; health care professional; caregiver; caregivers; technology; real-time; monitoring; aging; ageing

Introduction

Technological advances, principally computer-aided assessments and electronic health information exchanges, have enormous promise to promote healthy aging. Fortunately, older adults (ie, those aged ≥65 years) are open to using new technologies [1]. Thus, health care is adopting technologies that support real-time monitoring to promote personalized and preventative care [2].

Frailty is an age-related dynamic state characterized by deficits across multiple physiologic systems with increased vulnerability to stressors [3,4]. Consensus guidelines state that early detection of frailty is key to its management [5,6]. To this end, technology-based frailty tools have allowed health care professionals to screen, assess, identify, and develop care plans for frail persons [7-11]. Yet, physician-oriented frailty screening tools are resource intensive.

A patient-oriented frailty website is the next step in frailty care [12]. In Canada, approximately 23% of community-dwelling older adults are frail [13]. Although primary care frailty screening tools used by health care providers show promise in the identification and management of frailty [14], many community-dwelling older adults do not have the ability to undergo in-person frailty screening. Therefore, patient-oriented frailty screening tools in the form of websites and apps should be developed and implemented.

The concept of a patient-oriented frailty website is straightforward. Individuals go to a secure frailty website where they answer a series of health-related questions. The website, in turn, generates a frailty/health summary and score based on the user’s responses. In addition to a score, the website could also inform the user of their frailty status and whether they should seek medical counsel. Upon consent, users can share their responses and score with their care provider. In this capacity, a patient-oriented frailty website could identify those that are frail or at an increased risk of becoming frail. The website would be inclusive and encourage all older adults, regardless of health status, to participate. For healthy older adults that show no signs of frailty, the website would promote and reinforce their current healthy lifestyle and serve as a baseline. In this way, a frailty website is similar to the standard practice of well-baby visits—periodic examinations that screen for potential developmental health complications [15]. Therefore, well-baby visits and web-based frailty assessments are tools for the prevention and treatment of health issues that occur early and late in life, respectively.
To understand the interest in a patient-oriented frailty website, we conducted a planning project consisting of 2 components. In May 2023, we held a web-based educational symposium on frailty [16]. Experts in frailty spoke on the importance of frailty assessment and identification, approaches to preventing and mitigating frailty, frailty in primary care, and factors common to those who age gracefully beyond the age of 85 years. Recruitment for symposium audience and discussants was done via posters, which were distributed and published in local health authority newsletters, at academic institutions, and at a volunteer recruitment site for health research using REACH BC (the Ethical Considerations section provides more details). The audience was Canadian clinicians, researchers, caregivers, and community-dwelling older adults. Before and at the symposium, we invited interested audience members to participate in a discussion session scheduled 2 weeks later; individuals were excluded if they had not seen the symposium. Interested care providers, caregivers, and community-dwelling older adults returned to 1 of 2 discussion sessions. One discussion session included 7 caregivers and care providers, whereas the other session included 6 community-dwelling older adults; sample sizes were determined by best practices [17-19]. In contrast to a traditional qualitative research where findings are summarized then supplemented with participants’ quotes, in this viewpoint paper, we summarize the main findings of the discussion groups, highlighting the major barriers to developing a patient-oriented frailty website. To benefit other researchers developing similar health websites, we also provide potential solutions as observed throughout the literature and discuss caveats to those solutions, as summarized in Table 1.

<table>
<thead>
<tr>
<th>Barriers and solutions</th>
<th>Caveats</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completing a frailty assessment on the web is restrictive to frail persons.</td>
<td>Caregiver assists or completes assessment on frail person’s behalf.</td>
</tr>
<tr>
<td></td>
<td>Not all frail older adults will have a caregiver.</td>
</tr>
<tr>
<td></td>
<td>Requires a set of complex steps a frail person cannot do.</td>
</tr>
<tr>
<td>There are privacy and security concerns.</td>
<td>The website needs to be hosted from a reputable source with clear intent on its use.</td>
</tr>
<tr>
<td></td>
<td>It is unknown how users will perceive the trustworthiness of a frailty website.</td>
</tr>
<tr>
<td>Users may restrict researchers from health data analyses.</td>
<td>Implement federated learning.</td>
</tr>
<tr>
<td></td>
<td>Federated learning is still relatively novel and may be susceptible to attacks.</td>
</tr>
<tr>
<td>Users want to complete the assessment quickly.</td>
<td>Reduce the number of frailty questions.</td>
</tr>
<tr>
<td></td>
<td>Fewer questions may compromise the clinical usefulness of the assessment.</td>
</tr>
<tr>
<td>Health care providers need to know how to interpret and apply the information to care for frailty.</td>
<td>Develop models to treat frailty; implement training for health professionals.</td>
</tr>
<tr>
<td></td>
<td>This solution requires significant resources.</td>
</tr>
<tr>
<td>Being labeled as frail is counterproductive to combating frailty.</td>
<td>Educate users on frailty and reframe frailty in a more positive light (eg, well-being).</td>
</tr>
<tr>
<td></td>
<td>It is unknown how users will respond to their (potentially negative) scores.</td>
</tr>
<tr>
<td>Users want health resources tailored to them in addition to a frailty/health assessment.</td>
<td>Provide additional resources tailored to users’ responses.</td>
</tr>
<tr>
<td></td>
<td>Users may be deterred from seeking the expertise of medical professionals.</td>
</tr>
</tbody>
</table>

**Ethical Considerations**

This planning project received an exemption from the Fraser Health Research Ethics Board as it fell under quality improvement and evaluation studies. Stakeholders provided informed consent prior to the discussion session for review, following the Fraser Health guidelines. All stakeholders had the ability to opt out at any time for any reason and were reminded of this right prior to the start of the discussion session. Both discussion sessions were recorded for the purpose of transcription offline. Videos of the 2 discussions are stored on a secured drive that is both password protected and can only be accessed by approved Fraser Health employees. Stakeholders were compensated CAD $25 (US $18.27) for engaging in the symposium and 1 discussion session.

**A Website That Assesses Frailty Is Restrictive to the Frail**

The most substantial issue involving the implementation of a patient-oriented frailty website is its restrictiveness. A website targeting a subpopulation assumes that the targeted audience can access the website. On average, internet use among Canadian older adults is 68%, but it is only 41% for those aged ≥80 years [20]. In addition to age, health and frailty status are
also related to internet and computer use. For example, 73% of Canadian older adults in excellent or very good health use the internet, whereas 62% of older adults in fair or poor health use the internet [20]. Likewise, another group found that the frailest person was, the less likely they were to use a computer [21]. Thus, attempting to reach frail older adults through a website is a major hurdle.

A seemingly simple solution to this barrier is targeting caregivers (eg, partners, family members, friends, colleagues, or neighbors) and having caregivers assist or complete the questions on the frail person’s behalf. However, this solution only works for those with caregiver support. In one study based in the Netherlands, having a primary caregiver was reported in 32% of older adults who visited the emergency department [22]. This figure may be slightly lower in Canada. It is estimated between 23% [23] and 28% [24] of older Canadian adults have caregivers. Consequently, there is a chance that most frail community-dwelling older adults would not have assistance using a computer, navigating to the frailty website, and answering a series of health-related questions. Clearly, increasing caregiver support is one approach to increasing internet access among frail older adults.

Another solution to combat the restrictiveness of a frailty website is to also provide a paper version to frail older adults. Despite acknowledging the importance and benefits of digital screening tools, older adults have suggested that low-tech alternatives should also be available [25]. It would be convenient to make a paper pamphlet that has the same information and questions as a frailty website. This way, both paper and web-based frailty assessments would be interchangeable, allowing for conversion from one medium (ie, paper) to the other (ie, website) and vice versa. Pamphlets can be widely distributed and made available at family physicians’ and nurse practitioners’ offices, wellness and health clinics, pharmacies, and even mailed directly to older adults (eg, using Canada Post’s Precision Targeter). However, there are concerns regarding this approach as well. Assuming frail community-dwelling older adults obtained access to the frailty pamphlet survey, they would still need to fill it out (accurately), return it, schedule an appointment, and visit a health care professional. Some of these barriers can be mitigated by including a return address with free postage, using free door-to-door shared ride services, or scheduling a telehealth appointment. However, this process requires a series of complicated steps and older adults with mild cognitive impairment will be unable to complete them [26]. Other means of reaching frail persons are needed, but without significant assistance, it is unlikely that these individuals will be screened for proper frailty care.

**Older Adults Have Internet Privacy Concerns**

Despite a positive attitude toward web-based health services [27], older adults are concerned about privacy, especially when it is unclear by whom or how their medical data will be used [28]. For example, the probability of being identified was the single most important attribute when older adults considered internet privacy, even though sharing their medical data was viewed positively if it was to be used for science and the development of novel care and treatment [29]. The hesitancy of older adults to share their medical data reveals the need for clearly communicating how their information will be used—if they are to be comfortable consenting to its wider use. Similarly, a previous study found that older adults were more willing to share their data (eg, demographics, family relations, economic status, physical and mental illnesses, family history, medication, and health care service use) with family and hospitals compared to researchers or government agencies [30]. Public trust in the government is complex. For example, despite only 19% of Americans trusting the federal government, 70% viewed the US Food and Drug Administration and US Centers for Disease Control and Prevention favorably [31]. Thus, hosting a patient-oriented frailty website on a health authority or a particular branch of government, as opposed to a commercial site, with the clear intent to promote health, may increase the likelihood of the widespread adoption of a web-based frailty assessment tool among older adults. However, more work is needed to fully understand the actual adoption (as opposed to the hypothetical adoption) of a frailty website among older adults and how to securely transfer a user’s medical responses to a health authority without compromising privacy if the user consents.

**Users’ Restriction of Their Data May Limit Health Data Analyses**

The reluctance for individuals to share their health data for research purposes is understandable and may dictate the use of a frailty website. Researchers and health authorities may bemoan the decision for users to protect their data, which, in turn, may restrict our understanding of frailty. For example, knowing the frailty status and location of older adults who completed a web-based frailty assessment could reveal concentrated areas of frail older adults, suggesting environmental or societal risk factors and the need for additional services to be deployed to those locations. Realistically, this type of analysis can only be done with a user’s location and health data. Still, it is important to remember that the personal health data acquired from any website belongs to the user. It is at the discretion of the user (assuming they have the cognitive faculties to consent in the first place, a topic not discussed in our sessions) whether to share their data.

Hence the success of a frailty website is contingent upon the implementation of a privacy-first approach. A privacy-first approach underscores the users’ ownership of their data, using dynamic identifiers and storing data locally (ie, on a smartphone or computer), as opposed to a data-first approach, which prioritizes the retention and distribution of data, typically in a centralized location [32]. However, a happy medium exists where users can keep and protect their data while researchers and health authorities can advance frailty care through modern analysis.

New advances in data analysis [33] have been developed and used to underscore privacy-first approaches. Federated learning, for example, is a machine learning model that aggregates training results from multiple sources to create a consensus.
model without the need for data to leave a given device or system. A recent study found federated models achieved the same accuracy, precision, and generalizability as standard centralized statistical models using a variety of health data [34]. As a specific example, federated learning was used to predict treatment response in breast cancer patients using data behind a hospital’s firewall [35]. Emerging technologies and analyses have made it possible to have a patient-oriented frailty website that both ensures the privacy of the user and allows for analyses that will usher in better frailty care. A caveat to this is that federated learning is relatively new and will take time to implement across health authorities, and it may be prone to specific types of attacks [33].

**Users Want to Complete a Frailty Assessment Quickly**

A frailty website needs to be efficient and user-friendly [12] while adequately collecting health information that can assess frailty [3,4]. After the user completes a series of health-related questions on a frailty website, the website should produce a score (eg, “Your frailty score is 42/100, consider making an appointment with your doctor” or “Your biological age is 71, 6 years older than your actual age of 65”). One scoring approach could be a multisystem deficit-accumulation frailty index [36,37], which subscribes to the idea of an accumulation of deficits and is scored between 0 and 1, with 0 being no deficits present and 1 being all deficits present and fully expressed (in reliaiy, the score seldom exceeds 0.7, a limit of deficit accumulation); this approach has acceptable validity, reliability, and diagnostic test accuracy [38]. One benefit of a frailty score derived from a frailty index is that it can be interpreted by nonexperts. However, because the accumulation deficit model subscribes to the idea that frailty is a multisystem state, ideally a frailty website that adopts this approach would require 30 to 40 questions [39] across multiple domains (ie, physical, cognitive, psychological, and social). Research has shown that people are more likely to complete a survey if it takes 15 minutes or less [40]. In the context of a frailty website, completing 30 to 40 questions in 15 minutes would require users to spend 23 to 30 seconds on each question. Therefore, inaccurate responses and the respondent feeling rushed, frustrated, and stressed can be a concern.

In an attempt to make the website user-friendly, the number of questions would have to be reduced, potentially compromising its clinical utility [41]. A frailty index comprising fewer than 30 variables and questions can still be useful. In one study, researchers found that a frailty index constructed using 23 variables was just as accurate as one constructed using 70 variables [42]. An alternative approach to written multiple-choice questions may be a pictorial frailty assessment. The Pictorial Fit-Frail Scale [43] was recently developed for this purpose. It is fast (it took patients 6 minutes to complete) and comprehensive (the assessment covered 14 domains). However, agreement rates among Canadian and UK health care professionals were low (32% agreement for social, 44% agreement for mood, and 59% agreement for function), and it is unclear how patients understood each domain (averages were taken across 146 patients, caregivers, health care professionals, and general public participants) [43]. In contrast to the frailty index with precise grading of frailty, the phenotype model of frailty includes only 5 variables. Yet, the phenotype model requires grip strength and walking speed, measures that cannot be easily tested and may not be safe for many older adults to complete in their own homes. A methodological consensus regarding the definition of a frailty index, the variables that comprise it, and how they are scored has been encouraged [44-46].

**Frailty Score Needs to Be Interpretable**

In addition to a methodological consensus, output from a patient-oriented frailty website must be sufficient and interpretable for all health care professionals irrespective of location. For example, health care professionals across Canada would need to know what a frailty score of “42/100” means and how to prescribe the appropriate care. Our modern health care system is well designed for treating diseases but not for embracing the unique challenges of frailty among older adults, a population that experiences complex health issues [47].

Fortunately, the development of novel models [48] and guidelines [49] that address frailty in primary care is already underway. Moreover, frailty training would be required for physicians to prescribe appropriate frailty care. A systematic review found that there were limited frailty training programs for health professionals; however, the programs that did exist effectively increased frailty knowledge and competence in frailty assessments [50]. Although this would require significant resources, postponing or reducing frailty would result in a significant reduction in health care costs [51] and would pay for itself.

**Being Labeled Frail Is Counterproductive to Combating Frailty**

Numerous studies have agreed that there is a stigma around the concept of frailty and being labeled as frail. In one study, community-dwelling older adults reported that frailty was perceived as “approaching the end of their lives, malnourished and highly dependent on care” [52]. This same study also found that older adults are likely to reject the concept of frailty even when an objective measure may define them as frail [52]. Hence, a frailty website that labels a user frail may be counterproductive to the purpose of preventing and managing frailty.

In addition to education, another solution to combat the negative perceptions of frailty is to reframe the concept of frailty in a more positive light, using terminology such as “healthy aging” or “well-being” [52,53] or, within the context of a fit-frail score, “well-being score.” It should be noted, however, that positive output from a web-based frailty assessment (eg, a low fit-frail score, or a low biological age relative to chronological age) could also be counterproductive and have the same consequences as a negative output, deterring older adults from seeking medical care. Nevertheless, at least 1 study suggests otherwise. Among older adults who avoided medical care, 36% did so out of fear of a serious illness, and the likelihood of
avoidance was greater in those who self-reported a poorer health status [54]. Ultimately, it is unknown how users will respond to negative or positive outputs from a frailty website, and more research is needed to understand the potentially complicated reactions older adults will have after receiving their personalized health score and summary. Regardless, it is of the utmost importance that older adults are comfortable, feel safe, and are motivated (not discouraged) when using a frailty website.

**Frailty Websites Should Provide Additional Health Resources**

A frailty website should offer resources in addition to a health score and summary. A patient-oriented frailty website that gathers health information can be of great use for health care providers. It should also produce personalized resources based on the user's responses [55,56]. For example, if a community-dwelling older adult is deemed to be at risk of becoming frail and their deficits are primarily in a physical domain (ie, they are slightly overweight, sleep less than 6 hours a night, drink >6 alcoholic drinks a week, and have hypertension but are cognitively normal), the website ought to provide links to resources to meet those care needs. In this scenario, the website could suggest resources such as how to start and maintain an exercise program and provide best sleep hygiene practices in addition to recommending that the user make an appointment with their physician to manage hypertension.

Self-management is a potentially viable, low-cost approach to addressing frailty, but this is a relatively unexplored topic in frailty care [53]. Additionally, caution must be exercised with self-management. For instance, individuals that are provided resources may be lulled into a false sense of security and feel that their health can be managed without regular check-ups by the expertise of trained medical professionals. On the other hand, older adults may be unmotivated to complete a series of health-related questions if a website does not provide any immediate, clear feedback other than a score. There needs to be a happy medium where older adults are motivated to use a frailty website but also supplement the website with a medical expert's opinion. Future studies need to understand where that happy medium exists and incorporate it into frailty website development.

**Limitations**

This viewpoint paper is not without limitations. One such limitation is the lack of quotes from discussants. While quotes are typically included in qualitative research to strengthen findings, our viewpoint explicitly draws attention to barriers and potential solutions to the development of a patient-oriented frailty website. Further, traditional focus group studies that include theme discovery supplemented with quotes will be needed to better understand and address these barriers. Another limitation is the omission of technological barriers, which have been published elsewhere [12,57]. One important technological barrier is search engine optimization—a process of making modifications to a website to increase its visibility. Well-tuned site optimization strategies increase traffic to health websites, whereas poor search engine optimization can affect older adults’ experience in frailty assessment and health interpretation simply due to a missing hyperlink to the website. Finally, the barriers reported here may not generalize to populations outside of Canada. However, some of the barriers reported here also exist for older adults in Switzerland [36] and the United Kingdom [53,56]. Future studies should attempt to recruit samples representing broader geographical regions to promote equity in global health care.

**Conclusion**

A frailty website that can be used by the community to screen and identify older adults at an increased risk of frailty and health decline is an important step in geriatric care and public health. However, several barriers must be addressed in future research before the development of such websites. While some barriers have potential solutions, they come at a cost (eg, resources required for optimizing frailty models in primary care and patient-oriented frailty assessment training and support). Other potential solutions (eg, caregiver assistance and an accompanying paper-based frailty assessment) have their own challenges. Regardless, addressing these barriers, even partially, is a worthy goal. The early detection and management of frailty can lead to significant inroads to integrated care, benefiting the quality of life of older patients and their caregivers and the health of the aging population.

**Acknowledgments**

The authors sincerely acknowledge the distinguished speakers who generously shared their expertise during the educational symposium: Kenneth Rockwood, MD; John Muscedere, MD; Barry Clarke, MD; Grace Park, MD; Belinda Robinson, MD; and Angela Brooks-Wilson, PhD. The full recording of the symposium is available at the eFI-CGA (Electronic Frailty Index Based on the Comprehensive Geriatric Assessment) website [16]. We would also like to express our appreciation to the Fraser Health Authority Department of Evaluations and Research Services for research administrative support, the eFI-CGA IT support team for their valuable consultation in hosting the symposium materials on the website, and REACH BC for their assistance during the participant recruitment phase. We acknowledge the significant contributions of Alia Januwalla, BA, from the BC Support Unit for her knowledge translation inputs and support, and we extend our thanks to Hilary Low, BA, for administrative assistance. Finally, we would like to acknowledge all focus group participants for their time, commitment, and invaluable insights.

This planning project was supported by a planning and dissemination grant from the Canadian Institutes of Health Research (CIHR-PCS-185525). Additional funding for research operating and training came from a Canadian Institutes of Health Research project grant (CIHR PJT156210 eFI-CGA) and a Surrey Hospitals Foundation training award (FHG 2017-001).
Data Availability
Data sharing is not applicable to this article as this is a viewpoint article and represents the experiences and opinions of the authors.

Disclaimer
At no point was artificial intelligence used to aid the writing process.

Authors’ Contributions
BG contributed to conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, resources, supervision, and writing (original draft). SSC contributed to data curation, investigation, project administration, resources, and writing (review and editing). LG is a community patient advisor and contributed to conceptualization, investigation, project management, resources, supervision, and writing (review and editing). XS contributed to conceptualization, funding acquisition, team support, methodology, investigation, project management, resources, supervision, and writing (review and editing).

Conflicts of Interest
None declared.

References


Automatic Spontaneous Speech Analysis for the Detection of Cognitive Functional Decline in Older Adults: Multilanguage Cross-Sectional Study

Emilia Ambrosini1, PhD; Chiara Giangregorio1, MSc; Eugenio Lomurri1, MSc; Sara Moccia2, PhD; Marios Milis3, MSc; Christos Loizou4, PhD; Domenico Azzolini5, PhD; Matteo Cesari6, PhD; Manuel Cid Gala7, MSc; Carmen Galán de Isla7, MSc; Jonathan Gomez-Raja7, PhD; Nunzio Alberto Borghese8, PhD; Matteo Matteucci1, PhD; Simona Ferrante1,9, PhD

1Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milano, Italy
2BioRobotics Institute and Department of Excellence in Robotics and AI, Scuola Superiore Sant’Anna, Pisa, Italy
3SignalGeneriX Ltd, Limassol, Cyprus
4Department of Electrical Engineering, Computer Engineering and Informatics, Cyprus University of Technology, Limassol, Cyprus
5Geriatric Unit, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico Ca’ Granda Ospedale Maggiore Policlinico, Milano, Italy
6Ageing and Health Unit, Department of Maternal, Newborn, Child, Adolescent Health and Ageing, World Health Organization, Geneva, Switzerland
7Consejería de Sanidad y Servicios Sociales, Junta de Extremadura, Merida, Spain
8Department of Computer Science, University of Milan, Milano, Italy
9Laboratory of E-Health Technologies and Artificial Intelligence Research in Neurology, Joint Research Platform, Fondazione Istituto di Ricovero e Cura a Carattere Scientifico Istituto Neurologico Carlo Besta, Milano, Italy

Corresponding Author:
Emilia Ambrosini, PhD
Department of Electronics, Information and Bioengineering
Politecnico di Milano
Piazza Leonardo da Vinci 32
Milano, 20133
Italy
Phone: 39 0223999509
Email: emilia.ambrosini@polimi.it

Abstract

Background: The rise in life expectancy is associated with an increase in long-term and gradual cognitive decline. Treatment effectiveness is enhanced at the early stage of the disease. Therefore, there is a need to find low-cost and ecological solutions for mass screening of community-dwelling older adults.

Objective: This work aims to exploit automatic analysis of free speech to identify signs of cognitive function decline.

Methods: A sample of 266 participants older than 65 years were recruited in Italy and Spain and were divided into 3 groups according to their Mini-Mental Status Examination (MMSE) scores. People were asked to tell a story and describe a picture, and voice recordings were used to extract high-level features on different time scales automatically. Based on these features, machine learning algorithms were trained to solve binary and multiclass classification problems by using both mono- and cross-lingual approaches. The algorithms were enriched using Shapley Additive Explanations for model explainability.

Results: In the Italian data set, healthy participants (MMSE score\(\geq\)27) were automatically discriminated from participants with mildly impaired cognitive function (20\(\leq\)MMSE score\(\leq\)26) and from those with moderate to severe impairment of cognitive function (11\(\leq\)MMSE score\(\leq\)19) with accuracy of 80% and 86%, respectively. Slightly lower performance was achieved in the Spanish and multilanguage data sets.

Conclusions: This work proposes a transparent and unobtrusive assessment method, which might be included in a mobile app for large-scale monitoring of cognitive functionality in older adults. Voice is confirmed to be an important biomarker of cognitive decline due to its noninvasive and easily accessible nature.

(JMIR Aging 2024;7:e50537) doi:10.2196/50537
KEYWORDS

cognitive decline; speech processing; machine learning; multilanguage; Mini-Mental Status Examination

Introduction

According to “The 2021 Ageing Report by the European Commission,” life expectancy has shown a continuous trend over the past years [1]. As life expectancy increases, so does the number of people with dementia worldwide. Dementia is a neurodegenerative disease, which entails a long-term and gradual decrease in cognitive functionality, resulting in the reduction of patients’ autonomy and well-being, as well as worsening of the quality of life of their caregivers. The management of the increased number of older adults at risk of developing severe cognitive decline is a big challenge for health care systems, with the annual global cost expected to rise to US $2 trillion by 2030 [2]. These pathologies start silently up to 20 years before clear cognitive symptoms. However, there is increasing evidence that pharmaceutical interventions may be most effective at milder stages of dementia [3]. Thus, it is fundamental to find strategies that may anticipate the diagnosis [4-6].

Current diagnostic procedures require a thorough examination by medical specialists. The most employed tool for the first screening of cognitive function is the Mini-Mental Status Examination (MMSE). It is based on 30 questions that address short and long-term memory, attention span, concentration, language, and communication skills, as well as the ability to plan and understand instructions [7]. A score of 26 or higher is usually classified as normal. If the score is below 25, the result highlights a possible cognitive impairment, which may be classified as mild (21≤MMSE score≤26) or moderate to severe (MMSE score<20). Although this test has high sensitivity and specificity (87% and 82%, respectively) [8] and can be quickly administered, its employment is restricted within primary care facilities. Thus, faster, noninvasive, and automatic methods are needed to provide digital biomarkers for large-scale monitoring of cognitive functions in real-life scenarios [9].

In recent years, voice has been one of the most studied digital biomarkers since it allows cheap, noninvasive, ecological, rapid, and remote assessment of several aspects of a patient’s health status, such as the functionality of the respiratory system, cognitive decline, emotions, and heart dysfunctions [7,10-12]. Speech and language capacity is a well-established early indicator of cognitive deficits [13,14]. In the early phase of dementia, participants show alterations in the rhythm, resulting in a higher number of pauses, probably due to word-finding problems (ie, anomia and semantic paraphasia), worsening of verbal fluency [15-17], low speech rates, and decrease in the length of voiced segments [18-20]. Several studies have addressed the possibility of identifying signs of cognitive decline from voice recordings. Martínez-Sánchez and colleagues [21] analyzed the temporal parameters of reading fluency to discriminate between Spanish-speaking asymptomatic participants and those with Alzheimer disease (AD), and they were able to differentiate between patients with AD and healthy controls with an accuracy of 80% based on the speech rate. However, using a reading task introduces the possibility that participants’ fluency is affected by other factors such as educational level or visual impairment. Konig et al [22] demonstrated that it is possible to differentiate between dementia and mild cognitive impairment (MCI) in English-speaking participants based on voice features extracted from different tasks, for example, verbal fluency, picture description, counting down, and free speech, with a classification accuracy of 86%. Toth et al [23] showed that acoustic parameters such as speech rate, hesitation ratio, number of pauses, and articulation rate yield good results in discriminating between Hungarian-speaking participants with MCI and healthy controls. They analyzed a movie recall task and achieved an F1-score of 78.8%. Calzà et al [2] were able to discriminate between Italian-speaking healthy controls and participants with MCI by using random forest and support vector machine (SVM) with an F1-score of 75% by employing natural language processing. Finally, Bertini et al [24] achieved high performance (accuracy of 93% and F1-score of 88.5%) based on acoustic features extracted from spontaneous speech from a corpus of English-speaking participants, that is, Pitt Corpus, by using deep learning techniques. Nevertheless, natural language processing and deep learning require the analysis of raw data, thus having access to the recordings’ information content and endangering the participants’ privacy. Most previous works [2,21-24] aimed to distinguish participants with a proper diagnosis of AD or MCI from healthy participants. However, as far as we know, there are no studies investigating whether machine-learning algorithms based on voice features can identify early signs of functional cognitive decline detected by a decrease in the MMSE score.

In a previous study of our group [25], voice features automatically extracted from recordings of episodic storytelling could discriminate between Italian-speaking participants with normal cognitive functions (MMSE score≥27) and participants with mild cognitive decline (20≤MMSE score≤26) with an accuracy of 73%. Starting from this preliminary study, our study exploits acoustic features automatically extracted from spontaneous speech and machine learning techniques to support the early identification of cognitive function decline, meant as a reduction of the MMSE score. The main novelties involve the extension of a number of features, reduction of the computational time for feature extraction, and the multilanguage approach since both Spanish- and Italian-speaking participants were considered.

Methods

Participants and Data Collection

A sample of older adults were recruited in Italy (Lombardy region) and Spain (Extremadura region). In Italy, participants were recruited based on direct contact with the Geriatric Unit of the Foundation Scientific Institute for Research, Hospitalization and Healthcare (IRCCS) Ca’ Granda Ospedale Maggiore Policlinico (day hospital, ambulatory, and gymnasium). In Spain, people were recruited based on direct contact with professionals working in health care belonging to the Extremadura Health Ecosystem.
The essential requirement for participation was a good knowledge, at least oral, of the language of the country where the audios were recorded. Exclusion criteria were nonnative-speaking participants, clinically unstable participants, terminal illness (life expectancy <6 months), severe hearing or visual deficits, aphasia, and a score on the 30-item Geriatric Depression Scale >9. After providing informed consent to participate in the study, participants were met individually and they underwent the MMSE performed by health care professionals (geriatrician in Italy and neuropsychologist in Spain). Afterward, they were asked to tell 3 stories about their life for 2 minutes each without interruptions (positive, negative, and episodic) and to provide a 2-minute description of the “Cookie-Theft picture” of the Boston Diagnostic Aphasia Examination [26]. For each task, voice signals were recorded in separate .WAV files (16 kHz) by using an ad-hoc toolbox developed in MATLAB (MathWorks), through an external USB microphone. Participants were divided into 3 groups based on the MMSE score:

1. **Group 1**: MMSE score ≥27, that is, healthy participants
2. **Group 2**: 20 ≤ MMSE score ≤ 26, that is, participants with mild impairment of cognitive function
3. **Group 3**: 11 ≤ MMSE score ≤ 19, that is, participants with a moderate to severe impairment of cognitive function

The choice of the MMSE score for separation among the groups was employed since the aim was to detect the earliest symptoms of cognitive decline in the prediagnostic phase.

**Ethics Approval**

This study was approved by the ethics committee of Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico in Italy (ref: 1272018, approval date: March 15, 2018) and by the Comité Ético de Investigación Clínica de Badajoz in Spain (approval date: April 11, 2018).

**Feature Extraction and Statistical Analysis**

Data preprocessing and features extraction were performed employing an automatic algorithm implemented in MATLAB [25]. A positive speech polarity was imposed, and voice recordings were standardized. Afterward, the acoustic features described in Table 1 were extracted [20,21,23,27,28]. The features were grouped into 4 macrocategories according to their information content: voice periodicity, shimmer-related, syllabic, and spectral features. Feature extraction was repeated 3 times for voice segments lasting 5 seconds, 10 seconds, and 15 seconds to assess whether different time lengths can capture specific patterns. For each voice segment length, voice features extracted from the 4 audio recordings were substituted by their mean and standard deviation or their median and interquartile range, based on data set distribution, assessed by the Anderson-Darling normality test. Thus, each participant was represented by a single entry in the final data set, and 138 acoustic features (23 features × 3 segments length × 2 statistics) were computed for each entry. A 1-way analysis of variance for independent samples was applied to compare the 3 groups in terms of age. Due to their categorical nature, the Kruskal-Wallis test was applied to compare years of education and MMSE scores among groups. A Pearson chi-squared test was instead used for gender. Finally, generalized linear mixed models were defined in SPSS Statistics (version 28; IBM Corp) to evaluate whether acoustic features were significantly different among groups. Specifically, the mean (or median) values of the 23 acoustic features extracted from the 15-second segments were considered as the target for each model following a gamma regression distribution with a log link to the linear model. If significant differences were found, post hoc analysis with Bonferroni correction was also performed.
Table 1. Overview of the extracted features.

<table>
<thead>
<tr>
<th>Domain, feature description</th>
<th>Feature code</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Voice periodicity</strong></td>
<td></td>
</tr>
<tr>
<td>Unvoiced percentage, that is, percentage of aperiodic parts in the audio segment</td>
<td>F1</td>
</tr>
<tr>
<td>Duration of voiced and unvoiced segments, that is, mean, median, 15th and 85th percentiles of the parts of the signal with (voiced) and without (unvoiced) periodic nature</td>
<td>F2-F9</td>
</tr>
<tr>
<td>Percentage of voice breaks computed on the number of distances between consecutive pulses longer than 1.25 divided by the pitch floor (70 Hz) [27]</td>
<td>F10</td>
</tr>
<tr>
<td><strong>Shimmer</strong></td>
<td></td>
</tr>
<tr>
<td>Shimmer, that is, random cycle-to-cycle temporal changes of the amplitude of the vocal fold vibration [28]</td>
<td>F11</td>
</tr>
<tr>
<td><strong>Syllabic and pauses features</strong></td>
<td></td>
</tr>
<tr>
<td>Speech rate, that is, number of syllables per second [21]</td>
<td>F12</td>
</tr>
<tr>
<td>Percentage of phonation time, that is, the intrasyllabic and intersyllabic nuclei time &lt;250 ms divided by the total speech time [20,21]</td>
<td>F13</td>
</tr>
<tr>
<td>Articulation rate, that is, the number of syllables divided by the phonation time without pause [20,21]</td>
<td>F14</td>
</tr>
<tr>
<td>Mean duration of intersyllabic pauses &gt;250 ms [21]</td>
<td>F15</td>
</tr>
<tr>
<td>Mean duration of syllables [20,21]</td>
<td>F16</td>
</tr>
<tr>
<td>Number and mean duration of pauses of the audio segment [23]</td>
<td>F17-F18</td>
</tr>
<tr>
<td><strong>Spectral features</strong></td>
<td></td>
</tr>
<tr>
<td>Mean (SD) of pitch</td>
<td>F19-F20</td>
</tr>
<tr>
<td>Standard deviation of third formant (F3-SD)</td>
<td>F21</td>
</tr>
<tr>
<td>Speech temporal regularity, that is, temporal structure of the audio segment</td>
<td>F22</td>
</tr>
<tr>
<td>Centroid, that is, location of the center of mass of the spectral signal</td>
<td>F23</td>
</tr>
</tbody>
</table>

**Feature Selection and Classification**

Machine learning algorithms were trained to solve multiclass and binary classification problems (group 1 vs group 2 and group 1 vs group 3) starting from the extracted voice features, which were preliminary normalized.

**Classifiers**

SVM [29], logistic regression (LR), and CatBoost classifier (CAT) [30] were used. SVM is robust to noise in training data, since SVM decisions are only determined by the support vectors, while CAT represents the state of the art of boosting algorithms based on decision trees, and it has been proven to be very effective with small data sets with a high number of features. LR was investigated due to its simplicity and low computational cost. To achieve robust estimations despite the relatively small number of samples, the performance of each classifier was evaluated using stratified nested 10-fold cross-validation, which leads to the construction of an ensemble model via soft voting starting from each fold, obtaining a macromodel composed of 10 models trained on different subsets of data [31]. The classifier was selected according to the accuracy obtained in validation. Finally, a Kruskal-Wallis test was performed to determine whether there was a statistically significant difference between different classifiers in terms of accuracy.

**Parameter Setting**

Hyperparameter tuning was performed to limit overfitting with the nonlinear classifier. The following parameters were tuned for CAT through a randomized search method: bagging temperature, tree depth, l2 leaf regularization, and random strength. SVM was employed with a linear kernel and default parameters, and LR was also considered with default parameters. All the experiments were implemented using scikit-learn Python libraries, Catboost library, and Shapley Additive Explanation (SHAP).

**Feature Selection**

Due to the high dimensionality of the features set, the selection of the most informing features was performed through SHAP [32]. For each fold, starting from the entire set of features, the training was performed iteratively by computing the accuracy and the feature importance via SHAP for that specific iteration. At the end of each iteration, the 2 least significant features were removed until the minimum number of 6 features was reached. Therefore, the best model, that is, the one that achieved the best accuracy, was selected for each fold of the outer loop, and the model parameters were tuned for the identified set of features. As a result, 10 models trained on 10 different folds, each characterized by a different set of parameters and exploiting a different set of features, were obtained. The algorithm related to a single fold of the outer loop is summarized in Textbox 1. Finally, the ranking of the most informing features was implemented by summing up the unweighted mean of the Shapley values obtained at the end of the training of each fold for each feature.
Algorithm of feature elimination with Shapley Additive Explanations.

1: Train algorithm with whole set of features
2: Calculate model performance
3: Calculate feature importance with Shapley Additive Explanations
4: for feature in range (0, total features-6) do
   • Remove the k=2 least significant features
   • Train the model with the remaining features
   • Evaluate machine learning performance based on the scoring function
   • Calculate new features ranking with Shapley Additive Explanations explainer
5: end for
6: Best set is the one with the highest scoring function

Results

Characteristics of the Participants

Table 2 shows the characteristics of the recruited participants. A total of 266 participants were recruited: 133 Italian-speaking and 133 Spanish-speaking older adults. In the Italian data set, most participants in all groups were females. In contrast, in the Spanish data set, participants were balanced for gender in group 1 and unbalanced in favor of females in the other 2 groups. Overall, significant differences in terms of age (P=.03 and P=.001 for the Italian and Spanish data sets, respectively), MMSE scores (P=.001 for both data sets), and years of education (only for the Italian data set, P<.001) were found among the 3 groups, with people with severe impairment of the cognitive function being characterized by an older age in both data sets and by fewer years of education in the Italian data set.

Table 2. Characteristics of the participants.

<table>
<thead>
<tr>
<th></th>
<th>Group 1&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Group 2&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Group 3&lt;sup&gt;c&lt;/sup&gt;</th>
<th>P value</th>
<th>P value group 1 vs group 2</th>
<th>P value group 2 vs group 3</th>
<th>P value group 1 vs group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Italian data set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants, n</td>
<td>45</td>
<td>44</td>
<td>44</td>
<td>N/A&lt;sup&gt;d&lt;/sup&gt;</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>76.5 (4.9)</td>
<td>82.8 (4.6)</td>
<td>84.9 (5.7)</td>
<td>.03</td>
<td>.22</td>
<td>&gt;.99</td>
<td>.02</td>
</tr>
<tr>
<td>Gender (female/male)</td>
<td>39/6</td>
<td>33/11</td>
<td>37/7</td>
<td>.40</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MMSE (0-30), median (IQR)</td>
<td>30 (1)</td>
<td>24 (3)</td>
<td>16 (5)</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Years of education, median (IQR)</td>
<td>13 (3)</td>
<td>8 (8)</td>
<td>5 (5)</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>.36</td>
<td>&lt;.001</td>
</tr>
<tr>
<td><strong>Spanish data set</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Participants, n</td>
<td>43</td>
<td>45</td>
<td>45</td>
<td>N/A&lt;sup&gt;d&lt;/sup&gt;</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Age (years), mean (SD)</td>
<td>79.9 (7.5)</td>
<td>82.4 (6.9)</td>
<td>85.6 (6.6)</td>
<td>.001</td>
<td>.05</td>
<td>.27</td>
<td>.001</td>
</tr>
<tr>
<td>Gender (female/male)</td>
<td>21/22</td>
<td>36/9</td>
<td>27/18</td>
<td>.09</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MMSE (0-30), median (IQR)</td>
<td>28 (2)</td>
<td>23 (3)</td>
<td>17 (2)</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Years of education, median (IQR)</td>
<td>6 (5)</td>
<td>5 (4)</td>
<td>7 (4)</td>
<td>.25</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<sup>a</sup>Group 1: Mini-Mental Status Examination score≥27.
<sup>b</sup>Group 2: 20≤Mini-Mental Status Examination score<26.
<sup>c</sup>Group 3: 11≤Mini-Mental Status Examination score<19.
<sup>d</sup>N/A: not applicable.
<sup>e</sup>MMSE: Mini-Mental Status Examination.

Acoustic Feature Characteristics

Table 3 reports the results of the statistical analysis comparing acoustic features for the Italian and Spanish data sets. Voice periodicity features, particularly those related to unvoiced segments, were found to be significantly different among groups (P<.001 for mean, median, and 85th percentile of duration of unvoiced segments). Indeed, from group 1 up to group 3, a significant increase (P<.001 for mean and 85th percentile and P=.004 for median) in the unvoiced duration was found. Significant differences were found also for some syllabic features such as duration of pauses and syllables, which significantly increased with the decrease in the MMSE score.

Table 3. Acoustic feature characteristics.

https://aging.jmir.org/2024/1/e50537
as expected from literature [20]. The results of the statistical
analysis comparing acoustic features for the Italian and Spanish
data sets separately are reported in Multimedia Appendix 1
(Tables S1 and S2).
Table 3. Acoustic feature characteristics and significance between the 3 groups for the Italian and Spanish data sets.

<table>
<thead>
<tr>
<th>Domain, features</th>
<th>Group 1a (n=88)</th>
<th>Group 2b (n=88)</th>
<th>Group 3c (n=89)</th>
<th>P value</th>
<th>Group 1 vs group 2</th>
<th>Group 2 vs group 3</th>
<th>Group 1 vs group 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice periodicity, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unvoiced (%)</td>
<td>32.7 (10.4)</td>
<td>38.2 (13.5)</td>
<td>44.7 (11.8)</td>
<td>&lt;.001d</td>
<td>.004d</td>
<td>.003d</td>
<td>.001d</td>
</tr>
<tr>
<td>Duration of voiced segments (s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.08 (0.36)</td>
<td>1.03 (0.4)</td>
<td>0.98 (0.79)</td>
<td>.24</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Median</td>
<td>0.88 (0.32)</td>
<td>0.82 (0.35)</td>
<td>0.8 (0.75)</td>
<td>.33</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>15th percentile</td>
<td>0.26 (0.1)</td>
<td>0.24 (0.1)</td>
<td>0.22 (0.11)</td>
<td>.003d</td>
<td>.24</td>
<td>.29</td>
<td>.002d</td>
</tr>
<tr>
<td>85th percentile</td>
<td>2.03 (0.71)</td>
<td>1.93 (0.76)</td>
<td>1.77 (1.07)</td>
<td>.08</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Duration of unvoiced segments (s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>0.5 (0.14)</td>
<td>0.6 (0.2)</td>
<td>0.71 (0.23)</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
</tr>
<tr>
<td>Median</td>
<td>0.37 (0.12)</td>
<td>0.43 (0.13)</td>
<td>0.49 (0.17)</td>
<td>&lt;.001d</td>
<td>.001d</td>
<td>&lt;.001d</td>
<td>.04d</td>
</tr>
<tr>
<td>15th percentile</td>
<td>0.15 (0.02)</td>
<td>0.16 (0.03)</td>
<td>0.17 (0.03)</td>
<td>.08</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>85th percentile</td>
<td>0.91 (0.27)</td>
<td>1.14 (0.46)</td>
<td>1.42 (0.56)</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
</tr>
<tr>
<td>Voice breaks (%)</td>
<td>34.22 (10)</td>
<td>39.89 (13)</td>
<td>47.17 (11)</td>
<td>&lt;.001d</td>
<td>.001d</td>
<td>.001d</td>
<td>&lt;.001d</td>
</tr>
<tr>
<td>Shimmer, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shimmer</td>
<td>5 (0.56)</td>
<td>5.19 (0.67)</td>
<td>5.05 (0.75)</td>
<td>.17</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Syllabic and pauses features, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speech rate (syll/s)</td>
<td>3.92 (0.59)</td>
<td>3.52 (0.63)</td>
<td>3.78 (6.76)</td>
<td>.15</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Phonation (%)</td>
<td>70 (8)</td>
<td>64 (9)</td>
<td>64 (71)</td>
<td>.06</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Articulation rate (syll/s)</td>
<td>5.61 (0.44)</td>
<td>5.54 (0.5)</td>
<td>5.41 (0.46)</td>
<td>.02d</td>
<td>.89</td>
<td>.22</td>
<td>.01d</td>
</tr>
<tr>
<td>Mean intersyllabic duration (s)</td>
<td>0.14 (0.01)</td>
<td>0.15 (0.02)</td>
<td>0.16 (0.02)</td>
<td>&lt;.001d</td>
<td>.001d</td>
<td>.003d</td>
<td>&lt;.001d</td>
</tr>
<tr>
<td>Mean syllabic duration (s)</td>
<td>0.74 (0.17)</td>
<td>0.91 (0.29)</td>
<td>1.14 (0.37)</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
</tr>
<tr>
<td>Number of pauses</td>
<td>0.62 (0.19)</td>
<td>0.79 (0.31)</td>
<td>1.01 (0.48)</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
<td>&lt;.001d</td>
</tr>
<tr>
<td>Mean duration of pauses (s)</td>
<td>5.09 (1.26)</td>
<td>5.1 (1.42)</td>
<td>5.32 (1.23)</td>
<td>.51</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Spectral features, mean (SD)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>162 (25)</td>
<td>166 (26)</td>
<td>158 (25)</td>
<td>.11</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>SD</td>
<td>68 (12)</td>
<td>72 (12)</td>
<td>79 (15)</td>
<td>&lt;.001d</td>
<td>.10</td>
<td>.004d</td>
<td>&lt;.001d</td>
</tr>
<tr>
<td>F3-SD</td>
<td>466 (46)</td>
<td>490 (46)</td>
<td>484.92 (46)</td>
<td>.01d</td>
<td>.002d</td>
<td>&gt;.99</td>
<td>.02d</td>
</tr>
<tr>
<td>Speech temporal regularity</td>
<td>1749.5 (66)</td>
<td>1716.4 (67)</td>
<td>1687.8 (85)</td>
<td>&lt;.001d</td>
<td>.01d</td>
<td>.03d</td>
<td>&lt;.001d</td>
</tr>
<tr>
<td>Centroid</td>
<td>807.8 (154)</td>
<td>776.5 (165)</td>
<td>755.9 (193)</td>
<td>.12</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

aGroup 1: Mini-Mental Status Examination score≥27.
bGroup 2: 20≤Mini-Mental Status Examination score<26.
cGroup 3: 11≤Mini-Mental Status Examination score<19.
Multiclass Classification

Table 4 reports the results of the multiclass classification in terms of accuracy for 3 data sets: only Italian, only Spanish, and combination of Italian and Spanish participants. Overall, CAT achieved the best scores on the validation sets for the 3 data sets, but its performance considerably worsened when applied to the test sets. From the Kruskal-Wallis test, it can be seen that CAT achieved significantly better performance than LR for all data sets (P=.005, P=.02, and P=.03 for the Italian, Spanish, and Italian&Spanish data sets, respectively). A significant difference was also highlighted between SVM and CAT for the multilanguage data set (P=.01) and between SVM and LR for the Spanish data set (P=.003). Since there was no substantial difference in the accuracy between SVM and CAT, SVM was selected for its simplicity and further metrics, that is, receiver operating characteristic (ROC) curves, confusion matrices, and feature rankings are also reported (Figures 1-2). Overall, ROC curves (Figures 1A, 1C, and 1E) show a better trend for groups 1 and 3, whereas for group 2, the curve almost overlaps the bisector. The macro and micro averages of the areas under the curves achieved a fair score. The confusion matrices in Figures 1B, 1D, and 1F confirm this trend, with group 2 being the most misclassified in all 3 cases. For the Italian data set (Figure 1B), the model mainly misclassifies the participants from group 2 with those belonging to group 3, whereas for the Spanish data set (Figure 1D), participants from group 2 were mainly misclassified with participants from group 1.

Table 4. Classification accuracies on the validation and test sets for the multiclass classification among the 3 groups (healthy, mild, and severe symptoms) for the 3 data sets.

<table>
<thead>
<tr>
<th>Data set</th>
<th>CAT&lt;sup&gt;a&lt;/sup&gt;</th>
<th>SVM&lt;sup&gt;b&lt;/sup&gt;</th>
<th>LR&lt;sup&gt;c&lt;/sup&gt;</th>
<th>P value</th>
<th>CAT vs SVM</th>
<th>CAT vs LR</th>
<th>SVM vs LR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Italian, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.67 (0.03)</td>
<td>0.64 (0.02)</td>
<td>0.63 (0.03)</td>
<td>.006</td>
<td>.46</td>
<td>.005</td>
<td>.12</td>
</tr>
<tr>
<td>Test</td>
<td>0.54 (0.08)</td>
<td>0.57 (0.16)</td>
<td>0.59 (0.13)</td>
<td>N/A&lt;sup&gt;d&lt;/sup&gt;</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Spanish, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.63 (0.02)</td>
<td>0.64 (0.02)</td>
<td>0.60 (0.02)</td>
<td>.002</td>
<td>.88</td>
<td>.02</td>
<td>.003</td>
</tr>
<tr>
<td>Test</td>
<td>0.49 (0.09)</td>
<td>0.53 (0.11)</td>
<td>0.51 (0.15)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Italian&amp;Spanish, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.61 (0.01)</td>
<td>0.58 (0.02)</td>
<td>0.58 (0.02)</td>
<td>.008</td>
<td>.01</td>
<td>.03</td>
<td>.92</td>
</tr>
<tr>
<td>Test</td>
<td>0.53 (0.06)</td>
<td>0.54 (0.08)</td>
<td>0.52 (0.09)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

<sup>a</sup>CAT: CatBoost classifier.
<sup>b</sup>SVM: support vector machine.
<sup>c</sup>LR: logistic regression.
<sup>d</sup>N/A: not applicable.
Figure 1. Receiver operating characteristic curves and confusion matrices obtained with support vector machine for multiclass classification of the (A,B) Italian, (C,D) Spanish, and (E,F) Italian&Spanish data sets, respectively. (A,C,E): The dotted pink line corresponds to the microaveraged receiver operating characteristic curve, while the dotted blue curve corresponds to the macroaveraged one. (B,E,F): Labels 1, 2, and 3 on the x and y axes correspond to the group number.
Figure 2. Feature ranking for (A) Italian, (B) Spanish, and (C) Italian&Spanish data sets. Rank is represented from top to bottom from the most contributing to the least important feature.

**Binary Classification**

Tables 5 and 6 report the performance achieved for the binary classification, respectively, to distinguish group 1 (MMSE $\geq 27$) and group 2 ($20 \leq$ MMSE $\leq 26$) and group 1 and group 3 ($11 \leq$ MMSE $\leq 19$). SVM achieved the best scores on the validation sets compared to CAT and LR for the Italian&Spanish data sets in the discrimination between group 1 and group 2. However, the discrimination between group 1 and group 3 achieved a substantial equivalence among the 3 algorithms. As expected, better performance was obtained in the discrimination between healthy participants and those with severe impairment.

As for the multiclass scenario, the accuracy of the test sets worsened in all data sets, with the Spanish data set experiencing the largest decrease.

ROC curves, confusion matrices, and feature rankings for the Italian&Spanish data set were shown for SVM, which achieved the best performance, at least in distinguishing group 1 from group 2. Regarding ROC curves, the results were poor for the classification between healthy participants and those with mild impairment, with an area under the curve score of 0.65 (Figure 3A), as it can be noticed also by the confusion matrix in Figure 3B. Fair results were obtained for the ROC curve concerning the distinction between healthy participants and participants with impairment, with an area under the curve score of 0.77 (Figure 4A). Moreover, the confusion matrix (Figure 4B) shows a smaller number of misclassified participants. Feature rankings showed that the most informing features were mainly spectral features and features related to voice periodicity for the classification between healthy participants and participants with mild impairment (Figure 5A). In contrast, features related to syllables and pauses (in blue) were more important for classifying between healthy and older adults with severe impairment (Figure 5B). ROC curves, confusion matrices, and feature rankings related to binary classifications of the Italian&Spanish data sets are reported in Multimedia Appendix 1.
Table 5. Classification accuracies on the validation and test sets for the binary classification of group 1 (Mini-Mental State Examination score \( \geq 27 \)) versus group 2 (20 \( \leq \) Mini-Mental State Examination score \( \leq 26 \)).

<table>
<thead>
<tr>
<th>Data set</th>
<th>CAT(^a)</th>
<th>SVM(^b)</th>
<th>LR(^c)</th>
<th>(P) value</th>
<th>CAT vs SVM</th>
<th>CAT vs LR</th>
<th>SVM vs LR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Italian, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.80 (0.02)</td>
<td>0.84 (0.02)</td>
<td>0.79 (0.02)</td>
<td>(&lt;.001)</td>
<td>.007</td>
<td>.91</td>
<td>.02</td>
</tr>
<tr>
<td>Test</td>
<td>0.71 (0.14)</td>
<td>0.80 (0.14)</td>
<td>0.76 (0.16)</td>
<td>N/A(^d)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Spanish, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.74 (0.04)</td>
<td>0.79 (0.02)</td>
<td>0.76 (0.03)</td>
<td>.004</td>
<td>.005</td>
<td>.79</td>
<td>.03</td>
</tr>
<tr>
<td>Test</td>
<td>0.62 (0.15)</td>
<td>0.59 (0.16)</td>
<td>0.62 (0.17)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Italian &amp; Spanish, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.74 (0.02)</td>
<td>0.76 (0.01)</td>
<td>0.74 (0.02)</td>
<td>.06</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Test</td>
<td>0.64 (0.12)</td>
<td>0.65 (0.11)</td>
<td>0.65 (0.13)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

\(^a\) CAT: CatBoost classifier.  
\(^b\) SVM: support vector machine.  
\(^c\) LR: logistic regression.  
\(^d\) N/A: not applicable.

Table 6. Classification accuracies on the validation and test sets for the binary classification of group 1 (Mini-Mental State Examination score \( \geq 27 \)) versus group 3 (11 \( \leq \) Mini-Mental State Examination score \( \leq 19 \)) for the 3 data sets.

<table>
<thead>
<tr>
<th>Data set</th>
<th>CAT(^a)</th>
<th>SVM(^b)</th>
<th>LR(^c)</th>
<th>(P) value</th>
<th>CAT vs SVM</th>
<th>CAT vs LR</th>
<th>SVM vs LR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Italian, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.92 (0.02)</td>
<td>0.93 (0.02)</td>
<td>0.92 (0.02)</td>
<td>.38</td>
<td>N/A(^d)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Test</td>
<td>0.82 (0.14)</td>
<td>0.86 (0.18)</td>
<td>0.89 (0.14)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Spanish, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.84 (0.03)</td>
<td>0.83 (0.02)</td>
<td>0.82 (0.03)</td>
<td>.17</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Test</td>
<td>0.83 (0.12)</td>
<td>0.73 (0.11)</td>
<td>0.71 (0.15)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td><strong>Italian &amp; Spanish, mean (SD)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation</td>
<td>0.85 (0.02)</td>
<td>0.85 (0.01)</td>
<td>0.84 (0.01)</td>
<td>.05</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Test</td>
<td>0.79 (0.11)</td>
<td>0.78 (0.05)</td>
<td>0.81 (0.06)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

\(^a\) CAT: CatBoost classifier.  
\(^b\) SVM: support vector machine.  
\(^c\) LR: logistic regression.  
\(^d\) N/A: not applicable.
Figure 3. (A) Receiver operating characteristic curve and (B) confusion matrix for the binary classification of group 1 (Mini-Mental Status Examination score ≥27) and group 2 (20≤Mini-Mental Status Examination score ≤26) of the Italian&Spanish data set. AUC: area under the curve; ROC: receiver operating characteristic.

Figure 4. (A) Receiver operating characteristic curve and (B) confusion matrix for the binary classification between group 1 (Mini-Mental Status Examination score ≥27) and group 3 (11≤Mini-Mental Status Examination score ≤19) of the Italian&Spanish data set. AUC: area under the curve; ROC: receiver operating characteristic.

Figure 5. Feature ranking for the binary classifications of the Italian&Spanish data set. (A) Group 1 versus group 2; (B) group 1 versus group 3.
**Discussion**

An artificial intelligence–based classification pipeline has been implemented to evaluate the possibility of using voice analysis as a prescreening tool for detecting the impairment of cognitive function in a single and multilanguage approach. Multiclass and binary classification were performed on 3 data sets (Italian, Spanish, and a combination of Italian and Spanish data sets). For the multiclass tasks, the models obtained an accuracy of 57%, 53%, and 54% on the test set with SVM on the Italian, Spanish, and multilanguage data set, respectively. Regarding the binary classification, an accuracy of 80%, 59%, and 65% in the test set was achieved on the Italian, Spanish, and multilanguage data set, respectively, when distinguishing between healthy participants and those with the first symptoms of cognitive decline and an accuracy of 86%, 73%, and 78% for the classification between healthy participants and those with an MMSE score ≤ 19. The ROC curves in the multiclass task underlined how the participants with mild symptoms of cognitive decline are the most misclassified. This outcome aligns with expectations since participants belonging to this group exhibit mild impairment, indicated by an intermediate MMSE score. When having a deeper look into the misclassifications results (confusion matrices in Figure 1B and 1D), we observed that for the Italian data set, the model mainly misclassified participants from group 2 with those belonging to group 3, while for the Spanish data set, participants from group 2 were mainly misclassified with participants belonging to group 1. This result is in line with the results of the statistical analysis for the 2 data sets separately, which are reported in Multimedia Appendix 1 (Tables S1-S2): the statistical analysis highlighted a higher number of significantly different acoustic features between group 1 and group 2 for the Italian data set, while for the Spanish data set, there was a higher prevalence of acoustic features, which significantly differed between group 2 and group 3. A possible confounding factor might be the different distributions between the 2 data sets in terms of years of education, with a difference of 5 years between group 1 and group 2 for the Italian data set and a difference of only 1 year for the Spanish data set (Table 1).

The overall differences in the performance between the 2 languages may be explained by the heterogeneous demographic characteristics between the 2 data sets. Indeed, the distribution of participants in terms of gender, which highly affects acoustic features such as pitch [33], differed between the Italian and Spanish data set. In the Italian one, the distribution was more similar among the 3 groups, with a prevalence of females in each group, whereas for the Spanish data set, there was a prevalence of females in group 2 and group 3 compared to group 1, in which there was a balance between the 2 genders. Furthermore, the overall lower performance obtained on the Spanish data set may be related to the distribution of the MMSE scores among the groups. Indeed, there was a sharper separation among the 3 groups in the Italian data set, with a median MMSE score of 30 in group 1 and a median MMSE score of 24 in group 2, whereas the distribution of the scores in the Spanish data set was shrunk, with more participants being borderline among the groups (see Table 2).

The results highlighted that different sets of features are relevant depending on the considered language and the specific task. Indeed, shimmer was shown to be more relevant in Spanish-speaking participants, suggesting that an amplitude variation is predictive of a decline in cognitive function, whereas spectral features and those related to the voiced and unvoiced parts of speech were more important for predicting cognitive decline in Italian-speaking participants. The feature rankings of the classification tasks obtained with the multilanguage data set showed that the most informing features were a combination of those achieved for the 2 languages, when considered individually. This variability in the ranking of the features may be due to the change in prosody and accents of the languages themselves. Indeed, the Italian language is characterized by a wider spectral range compared with Spanish [34], which might explain why spectral features are predominant in the prediction of cognitive decline for the Italian data set. Nevertheless, these speculations need to be further explored in future studies.

Compared to that achieved by Calzà et al [2], we achieved slightly higher performance in the binary classification for distinguishing participants with mild cognitive decline from healthy participants when only the Italian data set was considered. Indeed, we achieved a test accuracy of 80%, while Calzà and colleagues [2] obtained an $F_1$-score of 75% on a manually checked corpus. However, there are several differences between these 2 studies. First, they considered not only acoustic features extracted from free speech but also lexical and syntactic features extracted with natural language processing as well as the demographic characteristics of the participants, such as age and years of education, which are considered important indicators of cognitive decline [35]. Conversely, we exploited only acoustic features automatically extracted from free speech, without considering any demographic features, to evaluate the possibility of exploiting this method for longitudinal monitoring. Moreover, in their work, Calzà and colleagues [2] recruited participants with a diagnosis of MCI based on a neuropsychological assessment, while in our work, we focused on the prescreening phase before an eventual diagnosis of MCI, and indeed, our groups were discriminated only based on the MMSE score. In another work [24], Bertini et al achieved instead higher performances, that is, 93% after data augmentation with a 20-fold cross-validation with acoustic features extracted from spontaneous speech from a corpus of English-speaking participants, that is, the Pitt Corpus, by applying deep learning techniques on a graphics processing unit. In their study, patients had a diagnosis of AD with the mean MMSE scores of the healthy control group of approximately 29 versus the AD group characterized by a mean score of 18. The lower performances of our model (accuracy of 86% to discriminate between group 1 and group 3 in the Italian data set) may be due to the use of the MMSE score only to distinguish between groups, which may have resulted in misclassification problems. Toth et al [23] achieved 75% accuracy from a binary classification task on Hungarian-speaking participants to distinguish healthy controls from those with MCI by using leave-one-out cross-validation on a set of 88 participants. Our results slightly outperformed their results on the Italian data set (accuracy of 80%), while we
achieved lower performances on the Spanish and the multilanguage data sets (accuracy of 62% and 65%, respectively). However, as in the previous studies [2,24], but differently from our study, Toth and colleagues [23] recruited participants with a diagnosis of MCI based on a neuropsychological assessment. Martínez-Sánchez et al [21] classified dementia among Spanish-speaking participants with an accuracy of 80%. The analysis was conducted to distinguish between 35 patients with AD and 35 healthy participants. They stated that fluency is an important aspect of cognitive decline from spontaneous speech, which was confirmed in our work by Figure 1D since the duration of syllables, phonation percentage, and articulation rate are in the top 5 most important features for the multiclass classification of the Spanish data set.

Our approach is based on acoustic features that can be automatically extracted on-the-fly on short speech segments. The satisfactory accuracy achieved with this approach to distinguish healthy participants from those with mild impairment (80% for the Italian data set) makes our results promising toward the design of a mobile app. Leveraging on this tool, an ecological and transparent mass screening of the early signs of cognitive decline can be performed, for example by analyzing free speech during phone calls. Moreover, since there is no need to store raw data and the information content of the speech is not exploited, this tool would preserve the speaker’s privacy.

This work has some limitations. The performance of the models on unseen data, that is, on the test set, worsened overall, probably due to the lack of the generalization power of the model; therefore, there is the need for larger data sets to have more robust classification models. Furthermore, regarding the Spanish data set, another limitation was the lower number of years of education of the recruited participants. Previous studies recruited only participants with more than 6 years of primary education to ensure that participants were fully literate [27]. Another limitation was the use of the MMSE score as the only method to allocate participants into different groups, without collecting information about an eventual diagnosis of MCI or AD, which might have brought to misclassification issues. Furthermore, as reported by Yancheva et al [36], MMSE is affected by a within-participant interrater standard deviation of 3.9 [37,38], which may have resulted in a further wrong group assignment for some participants. Finally, the MMSE test was administered by 2 different professional roles in the 2 recruitment sites—a geriatrician in Italy and a neuropsychologist in Spain, which may have introduced further differences between the 2 data sets.

This work confirmed that it is possible to detect early symptoms of cognitive function decline from the automatic analysis of acoustic features, exploiting a multilanguage approach. Overall, good performances by considering only acoustic features to discriminate between participants with different MMSE scores were achieved. The results obtained on the classification tasks are promising for the development of a screening tool for large-scale monitoring of cognitive function in community-dwelling older adults.

Acknowledgments

This work was supported by the European Project MOVECARE (Multiple-Actors Virtual Empathic Caregiver for the Elder; grant agreement: 732158) and the European project ESSENCE (Empathic Platform to Personally Monitor, Stimulate, Enrich, and Assist Elders and Children in their Environment; grant agreement: 101016112).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary results for the binary classification tasks of the Italian and Spanish data sets.

[PDF File (Adobe PDF File), 782 KB - aging_v7i1e50537_app1.pdf ]

References


Abbreviations

- **AD**: Alzheimer disease
- **CAT**: CatBoost classifier
- **IRCCS**: Scientific Institute for Research, Hospitalization and Healthcare
- **LR**: logistic regression
- **MCI**: mild cognitive impairment
- **MMSE**: Mini-Mental State Examination
- **ROC**: receiver operating characteristic
- **SHAP**: Shapley Additive Explanations
- **SVM**: support vector machine

**Edited by R Bjarnadottir; submitted 05.07.23; peer-reviewed by P Dabas, M Rouhizadeh, S Javorszky; comments to author 23.10.23; revised version received 18.12.23; accepted 12.02.24; published 29.04.24.**

Please cite as:


URL: https://aging.jmir.org/2024/1/e50537

doi:10.2196/50537

PMID:38386279
Assessing the Quality of ChatGPT Responses to Dementia Caregivers’ Questions: Qualitative Analysis

Alyssa Aguirre¹,²,*, MSW; Robin Hilsabeck³,*, PhD; Tawny Smith⁴, PharmD; Bo Xie⁵,⁶,*, PhD; Daqing He⁷,*, PhD; Zhendong Wang⁷,*, MS; Ning Zou⁷,*, MS

* all authors contributed equally

Corresponding Author:
Alyssa Aguirre, MSW

Abstract

Background: Artificial intelligence (AI) such as ChatGPT by OpenAI holds great promise to improve the quality of life of patients with dementia and their caregivers by providing high-quality responses to their questions about typical dementia behaviors. So far, however, evidence on the quality of such ChatGPT responses is limited. A few recent publications have investigated the quality of ChatGPT responses in other health conditions. Our study is the first to assess ChatGPT using real-world questions asked by dementia caregivers themselves.

Objectives: This pilot study examines the potential of ChatGPT-3.5 to provide high-quality information that may enhance dementia care and patient-caregiver education.

Methods: Our interprofessional team used a formal rating scale (scoring range: 0-5; the higher the score, the better the quality) to evaluate ChatGPT responses to real-world questions posed by dementia caregivers. We selected 60 posts by dementia caregivers from Reddit, a popular social media platform. These posts were verified by 3 interdisciplinary dementia clinicians as representing dementia caregivers’ desire for information in the areas of memory loss and confusion, aggression, and driving. Word count for posts in the memory loss and confusion category ranged from 71 to 531 (mean 218; median 188), aggression posts ranged from 58 to 602 words (mean 254; median 200), and driving posts ranged from 93 to 550 words (mean 272; median 276).

Results: ChatGPT’s response quality scores ranged from 3 to 5. Of the 60 responses, 26 (43%) received 5 points, 21 (35%) received 4 points, and 13 (22%) received 3 points, suggesting high quality. ChatGPT obtained consistently high scores in synthesizing information to provide follow-up recommendations (n=58, 96%), with the lowest scores in the area of comprehensiveness (n=38, 63%).

Conclusions: ChatGPT provided high-quality responses to complex questions posted by dementia caregivers, but it did have limitations. ChatGPT was unable to anticipate future problems that a human professional might recognize and address in a clinical encounter. At other times, ChatGPT recommended a strategy that the caregiver had already explicitly tried. This pilot study indicates the potential of AI to provide high-quality information to enhance dementia care and patient-caregiver education in tandem with information provided by licensed health care professionals. Evaluating the quality of responses is necessary to ensure that caregivers can make informed decisions. ChatGPT has the potential to transform health care practice by shaping how caregivers receive health information.

(JMIR Aging 2024;7:e53019) doi:10.2196/53019

KEYWORDS
Alzheimer’s disease; information technology; social media; neurology; dementia; Alzheimer disease; caregiver; ChatGPT

Introduction

Older adults have responded to the COVID-19 pandemic by expanding their medical care to the use of web-based platforms [1]. Indeed the internet has become the most common source of information among dementia caregivers [2], and with recent advances in artificial intelligence (AI), caregivers will
increasingly use AI to obtain information about health [3,4]. ChatGPT by OpenAI [5], an innovative, dialogue-based large language model that responds to complex natural language inquiries, holds great promise to improve the quality of life of patients with dementia and their caregivers by providing high-quality responses to meet their needs for information [4]. On the other hand, several studies have highlighted the limitations of generative AI models in health care, citing the lack of trust and reliability as some of the primary challenges [6,7]. Although there have been studies on the quality of ChatGPT responses to common questions about heart disease [8], cirrhosis [9], and bariatric surgery [10], to our knowledge, no studies have examined the quality of ChatGPT responses to real-world questions posed by dementia caregivers. We have addressed this gap by examining the quality of ChatGPT-3.5 responses to complex questions posted by dementia caregivers on social media.

**Methods**

**Overview**

From January to May 2023, a total of 60 social media posts representing dementia caregivers’ needs for information in 3 areas (memory loss and confusion, aggression, and driving; 20 posts per area) were selected from Reddit, a popular social media platform. These topics were chosen because they are common clinical themes that are often complex and difficult to navigate with potential safety implications. Four seed posts were used in each area to discover the additional 16 posts. Posts were excluded if the poster’s main question did not fall into the 3 aforementioned areas as verified by dementia clinicians or if the poster declared they were “venting” and/or no specific question was asked. Posts that were unclear on whether the person had a dementia diagnosis were excluded to avoid assessing posts that were not clearly dementia related. Word count for posts in the memory loss and confusion category ranged from 71 to 531 (mean 218; median 188), aggression posts ranged from 58 to 602 (mean 254; median 200), and driving posts ranged from 93 to 550 (mean 272; median 276). Of the 60 posts, the caregiver described the person with dementia as their parent (n=34, 56%), grandparent (n=22, 36%), uncle (n=2, 3%), or spouse (n=1, 1.6%). One post did not report relationship. The gender of the person with dementia was described as female in 57% (n=34) of posts and as male in 42% (n=25) of posts. One post did not report gender.

Three clinicians, each having more than 15 years of experience with patients with dementia and their caregivers, but from diverse disciplines (pharmacy, neuropsychology, and social work), assessed ChatGPT responses to the 60 posts using an adapted rating scale based on Hurtz et al’s [11] levels of cognitive complexity pertaining to clinical decision-making (Table 1). Responses received 1 point for each of the following characteristics: **factuality**, **interpretation**, **application**, **synthesis**, and **comprehensiveness**, with a scoring range of 0-5 for each response, where higher scores indicate higher quality. Table S1 in Multimedia Appendix 1 presents examples of posts for each topic area, ChatGPT responses, and clinician ratings for each response category.

**Table . Description of rating scale categories used to measure the quality of ChatGPT responses.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factuality</td>
<td>Response did not contain inaccurate or false information.</td>
</tr>
<tr>
<td>Interpretation</td>
<td>Response adequately interpreted the poster’s main need, correctly disregarded nonpriority details, and did not recommend strategies that the poster had already tried.</td>
</tr>
<tr>
<td>Application</td>
<td>Response suggested tangible actions (eg, educational information, a change the caregiver could make, and communication strategies such as validation and redirection).</td>
</tr>
<tr>
<td>Synthesis</td>
<td>Response contained follow-up recommendations as needed (referrals to help beyond the caregiver-patient dyad, such as support groups, health care professionals, or other community resources).</td>
</tr>
<tr>
<td>Comprehensiveness</td>
<td>Response had strong depth, breadth; response was thorough and complete.</td>
</tr>
</tbody>
</table>

Although the results reported in this paper were based on raters’ consensus scores, we acknowledge the potential benefits of expanding on ChatGPT responses that originally received different scores. Initially, 1 rater gave a point for comprehensiveness when the majority of suggestions they would provide clinically were conveyed in ChatGPT’s response, but another rater did not give the point if they felt it was missing anything at all. It was agreed upon during consensus that if the majority of recommendations were provided, ChatGPT responses would receive full credit for **comprehensiveness**.

**Ethical Considerations**

This study was approved by the institutional review boards of The University of Texas at Austin (STUDY00003358) and the University of Pittsburgh (STUDY200200007).

**Results**

ChatGPT responses in the memory loss and confusion category ranged from 89 to 276 words (mean 170; median 165), 91 to 372 words in the aggression category (mean 221; median 234), and 65 to 359 words in the driving category (mean 175; median 130). At least 2 clinicians agreed on the ratings for all ChatGPT responses, with any disagreements resolved by discussion.

https://aging.jmir.org/2024/1/e53019
ChatGPT scores ranged from 3 to 5. Overall, of the 60 responses, 26 (43%) received 5 points, 21 (35%) received 4 points, and 13 (21.7%) received 3 points (Table 2), suggesting high quality. There were no responses that scored a 0, 1, or 2; there were no fabricated responses; and no responses were considered harmful to posters. ChatGPT received the lowest ratings in comprehensiveness, followed by interpretation, and the highest ratings in synthesis, with only 2 out of 60 posts failing to receive the point (Table 3).

### Table 1. Number of ChatGPT points for each topic.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Memory loss and confusion (n=20), n (%)</th>
<th>Aggression (n=20), n (%)</th>
<th>Driving (n=20), n (%)</th>
<th>Total (N=60), n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factuality</td>
<td>17 (85)</td>
<td>19 (95)</td>
<td>20 (100)</td>
<td>56 (93)</td>
</tr>
<tr>
<td>Interpretation</td>
<td>17 (85)</td>
<td>17 (85)</td>
<td>13 (65)</td>
<td>47 (78)</td>
</tr>
<tr>
<td>Application</td>
<td>20 (100)</td>
<td>17 (85)</td>
<td>17 (85)</td>
<td>54 (90)</td>
</tr>
<tr>
<td>Synthesis</td>
<td>18 (90)</td>
<td>20 (100)</td>
<td>20 (100)</td>
<td>58 (96)</td>
</tr>
<tr>
<td>Comprehensiveness</td>
<td>9 (45)</td>
<td>15 (75)</td>
<td>14 (70)</td>
<td>38 (63)</td>
</tr>
</tbody>
</table>

### Discussion

#### Principal Findings

In this study, ChatGPT responses to complex, real-world questions posted by dementia caregivers were assessed by dementia clinicians using a clinical decision-making rating scale. ChatGPT was found to produce high-quality responses, suggesting the potential of online chatbots to be a useful source of health information for dementia caregivers. The majority of responses contained factual information (n=56, 93%), with 78% (n=47) of responses correctly interpreting the poster’s main need. The majority (n=54, 90%) of ChatGPT responses contained tangible actions the caregiver could apply to their situation. In only 2 instances, follow-up referrals were not suggested when reviewers felt recommendations were needed.

ChatGPT also had limitations, primarily in the areas of interpretation and comprehensiveness. In 22% (n=13) of posts, ChatGPT recommended strategies that posters had already explicitly tried, or missed subtleties that affected the accuracy of recommendations, such as failing to recognize that a person placed in a “home” meant a nursing home facility and not a traditional home. In another instance, ChatGPT recommended considering short-term hospitalization, but the poster already disclosed the person with dementia was currently hospitalized. In 37% (n=22) of posts, ChatGPT’s response did not include information that dementia clinicians felt was important or was unable to anticipate future problems that a human clinician might choose to address in response to the same post. For example, if ChatGPT recommended a driving test, it did not suggest what to do if the patient in question refused to take the driving test. The data suggest that ChatGPT has strengths in providing objectively correct information (factuality, application, and synthesis) but is less successful in contextualizing the information it provides (interpretation and comprehensiveness).

#### Limitations

Study limitations included potential sample bias and small sample size. Very few posters in this study identified as a spousal caregiver (n=1, 1.6%) even though national studies report that 60% of dementia caregivers are a spouse or partner [12]. In selecting social media posts for inclusion, we included only those in which it was clear that the individual had a diagnosis of dementia. Historically, racial and ethnic minority groups are less likely to seek or receive a dementia diagnosis; thus, our sample may have been skewed for race and ethnicity. Posts were from one specific platform, which risked including caregivers with a certain level of technology access and literacy. This study did not evaluate differences in ChatGPT responses at multiple time points, so no conclusions can be made regarding reproducibility. Raters were aware that responses were generated by ChatGPT, which could have influenced stricter grading. Although our 5-point scale graded specific aspects of ChatGPT responses, it might have had a ceiling effect.

#### Conclusions

This study contributes to the currently small but rapidly growing literature on AI’s potential to assist patient-caregiver education by providing high-quality information. Our study illustrates that ChatGPT-3.5 can provide high-quality responses to most questions in the areas of memory loss and confusion, aggression, and driving. Future research should examine family caregivers’ receptiveness to using ChatGPT, as well as the usefulness of the responses from the perspective of family caregivers. Validated rating scales to assess the quality of ChatGPT responses are still in progress; the field would benefit from a
reliable, validated method to evaluate the quality of AI responses to health care questions. We encourage future studies to expand on our findings and investigate how ChatGPT might be used in tandem with information provided by licensed health care professionals.

Acknowledgments
This work was supported in part by the National Institute on Aging of the National Institutes of Health (award number R56AG075770). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Editorial support with manuscript development was provided by Dr John Bellquist at the Cain Center for Nursing Research at The University of Texas at Austin School of Nursing.

Conflicts of Interest
TS was employed by The University of Texas at Austin during the submission of this work but has since changed positions and is employed by Otsuka America Pharmaceutical, Inc.

Multimedia Appendix 1
Scoring of responses generated from ChatGPT.

References

Abbreviations
- AI: artificial intelligence
Positive Emotional Responses to Socially Assistive Robots in People With Dementia: Pilot Study

Eri Otaka, MD, PhD; Aiko Osawa, MD, PhD; Kenji Kato, PhD; Yota Obayashi, OT, PhD; Shintaro Uehara, PT, PhD; Masaki Kamiya, OT, MS; Katsuhito Mizuno, MD, PhD; Shusei Hashide, MD; Izumi Kondo, MD, PhD

Abstract

Background: Interventions and care that can evoke positive emotions and reduce apathy or agitation are important for people with dementia. In recent years, socially assistive robots used for better dementia care have been found to be feasible. However, the immediate responses of people with dementia when they are given multiple sensory modalities from socially assistive robots have not yet been sufficiently elucidated.

Objective: This study aimed to quantitatively examine the immediate emotional responses of people with dementia to stimuli presented by socially assistive robots using facial expression analysis in order to determine whether they elicited positive emotions.

Methods: This pilot study adopted a single-arm interventional design. Socially assistive robots were presented to nursing home residents in a three-step procedure: (1) the robot was placed in front of participants (visual stimulus), (2) the robot was manipulated to produce sound (visual and auditory stimuli), and (3) participants held the robot in their hands (visual, auditory, and tactile stimuli). Expression intensity values for “happy,” “sad,” “angry,” “surprised,” “scared,” and “disgusted” were calculated continuously using facial expression analysis with FaceReader. Additionally, self-reported feelings were assessed using a 5-point Likert scale. In addition to the comparison between the subjective and objective emotional assessments, expression intensity values were compared across the aforementioned 3 stimuli patterns within each session. Finally, the expression intensity value for “happy” was compared between the different types of robots.

Results: A total of 29 participants (mean age 88.7, SD 6.2 years; n=27 female; Japanese version of Mini-Mental State Examination mean score 18.2, SD 5.1) were recruited. The expression intensity value for “happy” was the largest in both the subjective and objective assessments and increased significantly when all sensory modalities (visual, auditory, and tactile) were presented (median expression intensity 0.21, IQR 0.09-0.35) compared to the other 2 patterns (visual alone: median expression intensity 0.10, IQR 0.03-0.22; P<.001; visual and auditory: median expression intensity 0.10, IQR 0.04-0.23; P<.001). The comparison of different types of robots revealed a significant increase when all stimuli were presented by doll-type and animal-type robots, but not humanoid-type robots.

Conclusions: By quantifying the emotional responses of people with dementia, this study highlighted that socially assistive robots may be more effective in eliciting positive emotions when multiple sensory stimuli, including tactile stimuli, are involved. More studies, including randomized controlled trials, are required to further explore the effectiveness of using socially assistive robots in dementia care.

Trial Registration: UMIN Clinical Trials Registry UMIN000046256; https://tinyurl.com/yw37auan

(JMIR Aging 2024;7:e52443) doi:10.2196/52443

KEYWORDS
dementia care; robotics; emotion; facial expression; expression intensity; long-term care; sensory modality; gerontology; gerontechnology
Introduction

The number of people with dementia is increasing globally, and it is estimated that it will reach 152 million by 2050 [1]. The provision of adequate social care for people with dementia is a major public health concern in many countries. The neurodegenerative nature of dementia affects memory, cognitive function, and more, resulting in a range of noncognitive symptoms, including changes in behavior, emotion, and social functioning. The most frequent changes include agitation, depression, and apathy. These emotional changes are reported as the most challenging aspect of dementia care by many caregivers [2]. On the other hand, positive emotions such as joy or comfort are relatively preserved until the terminal stage of the disease [3,4]. Previous studies suggest that the arousal of positive emotions may enhance cognitive function, presumably through amygdala activation; therefore, stimuli with a positive valence may enhance the effect of the rehabilitative approach for patients with dementia [3,5]. Considering these facts, interventions and caregiver involvement that can evoke positive emotions and suppress negative psychological responses are important strategies that should be actively implemented in long-term care for people with dementia to maintain residual functions and alleviate the burden of care.

As examples of emotion-related interventions, music therapy [6,7] and occupational therapy [8,9] have been shown to be effective in terms of emotional control. However, due to the rapid increase in the number of people with dementia and the shortage of dementia caregivers [1,10,11], there is a lack of staffing power to provide such nondrug therapies broadly and equally. In recent years, clinical applications of socially assistive robots have been used to provide high-quality emotional support and companionship [12-14]. Socially assistive robots are machines designed to provide assistance in the caregiving process through social rather than physical means and are equipped with a social interface to enable interaction with the user [15-17]. One systematic review and meta-analysis, as well as one scoping review, found that Paro, a baby seal-shaped socially assistive robot, has significant effects on agitation and depression [12,18], while another systematic review and meta-analysis concluded that there is little evidence that people with dementia derive benefits from socially assistive robots for cognition or neuropsychiatric symptoms when considering various types of robots, although they are feasible and acceptable [11]. On the other hand, one small between-groups comparison study reported that a certain type of socially assistive robot showed a negative effect in participants with cognitive decline, based on an examination of immediate neurophysiological changes [19].

When using socially assistive robots in clinical practice, one important aspect to consider is the immediate response of persons with dementia. In general, when confronted with a new robot or technology, a relatively positive immediate response known as the novelty effect [20] tends to be observed. In contrast, in people with cognitive decline, the immediate response to robots is reported to be somewhat stressful rather than positive [19]. These findings suggest that people with dementia, or those with memory and other cognitive impairments, may have a different immediate response compared to the general public. For example, people with dementia have a reduced ability to process multiple sensory stimuli [21]; therefore, they might have difficulty accepting and integrating multiple unfamiliar stimuli (eg, shapes, lights, sounds, and touch) provided simultaneously by the robot in the first interaction. Since they are prone to mental stress when they do not understand a situation [22], these stimuli from the robots could cause a tense or negative response. Nonetheless, no studies have verified how socially assistive robots are perceived by people with dementia from the perspective of having to process multiple sensory modalities. Moreover, an immediate response from the person with dementia is crucial in clinical settings because it helps care providers confirm the effectiveness of the robot on the spot and make precise decisions about whether to continue using the robot. Therefore, it is beneficial for clinical applications to focus on understanding the immediate responses of people with dementia when they are given multiple sensory modalities from socially assistive robots.

For an objective and better understanding of these issues, the signs of emotional responses should be quantified using appropriate techniques. Given that verbal skills tend to be impaired in people with dementia [23,24], it is important to use not only self-reported outcomes but also objective measures that can be obtained with minimal burden. For example, in the field of psychology, facial expression is considered a differentiated indicator of inner emotions [25,26]. According to recent reports, analysis of facial imaging using facial expression analysis software is able to quantify facial expressions and estimate emotions with good validity [27,28]. Applying these technologies to the investigation of the use of socially assistive robots among people with dementia will enable detailed and empirical verification of their effects, such as responses to the different sensory stimuli mentioned above, which are difficult to detect with subjective scales.

In this context, this study aimed to quantitatively evaluate the psychological and emotional reactions evoked in people with dementia to stimuli derived from socially assistive robots using facial expression analysis of facial video clips. In particular, we investigated how immediate responses changed as the modalities of sensory stimulation provided by the robot increased. Furthermore, from the perspective of eliciting positive emotions, which are beneficial to dementia care, this study also examined the differences in the emotion of joy elicited by the different types of robots.

Methods

Ethical Considerations

The study protocol was approved by the institutional ethics committee of the National Center for Geriatrics and Gerontology (1539) and prospectively registered in the UMIN Clinical Trial Registry (UMIN000046256). All participants with a Clinical Dementia Rating (CDR) scale [29] score of 0 or 1 provided informed consent themselves in accordance with the Declaration of Helsinki. For those who were considered to have an insufficient capacity to consent due to cognitive decline equivalent to CDR 2 and 3, informed consent was obtained from
their family members, and the procedures were explained to the participants in plain language to obtain their approval. To comply with ethical principles, all data collected were anonymized and stored in a locked file or on a password-protected computer.

**Study Design and Setting**

This pilot study was conducted as a single-arm, self-controlled, interventional study. Two local nursing care facilities that had no previous experience implementing socially assistive robots were selected as the experiment sites.

**Participants**

Participants were recruited among the residents of the 2 nursing homes. The inclusion criteria were as follows: a significant decline in cognitive function interfering with independence in the performance of everyday activities, the ability to maintain a sitting position for 15 minutes or more, the ability to communicate using simple words, and the ability to follow 2-step instructions. These criteria were first assessed by nursing home staff members through assessments performed as part of daily nursing care procedures. Regarding cognitive decline, candidates either had a previous formal diagnosis of dementia from their physician or received a diagnosis from one of the researchers (EO, a physician); they also had confirmed evidence that cognitive decline was present and that the decline was not due to delirium or other mental disorders. Ultimately, all the participants met the diagnostic criteria for dementia in the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition* (*DSM-5*). Those with unstable physical or mental conditions or evident higher cognitive dysfunction due to causes other than dementia were excluded.

**Procedures**

The participants were taken to a private room or a place with minimum environmental noise, and their faces were recorded in a resting state for 30 seconds. This was referenced as the control image for calibrating the facial expression analysis (to be described below). Next, the socially assistive robots were presented to the participant by a familiar staff member in 1 session using a predetermined 3-step procedure. The staff explained in advance that they wanted the participants to share how they felt after experiencing the robots. In the first step, the robot was placed on a desk in front of the participant (visual stimulus). In the second step, the robot was manipulated to produce a gentle voice or meow (visual and auditory stimuli). In the third step, the participant was encouraged to touch the robot (visual, auditory, and tactile stimuli) and was able to handle it freely, including petting and holding. The 3 patterns of sensory stimulus produced by the robots were presented for approximately 30 seconds in the context of assessing the participant’s immediate responses unless the participant refused (Figure 1).
Each participant was presented with 3 types of commercially available robot with typical characteristics—a humanoid-type robot capable of voice communication (RoBoHoN; Sharp Corp), a doll-type robot with the appearance of a stuffed toy and a voice recognition and reproduction system (Chapit; RayTron Inc), and an animal-type (cat-shaped) robot that can meow, move its tail, and recognize sound (Amaenbou-Nekochan; Digirec Co, Ltd)—resulting in 3 sessions per participant. The order of presentation, which was determined in advance using a random number table, differed for each participant (Figure 2).
Assessments of Emotional Responses

Subjective Assessment
At the end of each robot presentation session, the degree of subjective feelings (happy, sad, angry, surprised, scared, and disgusted) was recorded using a 5-point Likert scale (1: not at all; 2: very little; 3: so-so; 4: somewhat; 5: very much). The participants were shown a scale with words corresponding to each of the 5 points of the scale and were asked to select the point that best described their current feelings. This was assessed a total of 3 times (ie, once at the end of each robot presentation).

Video Recording and Facial Expression Analysis
For the study procedures, a video camera (Logicool StreamCam; Logitech Inc) was configured on a desk to capture the participants’ faces from the front. In case this camera’s view was blocked by arm movements or the robot, another camera (Handycam HDR-CX470, Sony Corp) was fixed on a tripod positioned 30 degrees diagonally in front of the participant.

The videos were recorded continuously during the session in full HD (1920 x 1080 pixels) at 60 frames per second. To optimize the sensitivity and accuracy of the facial expression analysis, we segmented the video clips such that each clip contained 1 sensory stimulation pattern (visual only; visual and
auditory; or visual, auditory, and tactile), and then cropped them to a suitable resolution that focused on the facial region. If necessary, the brightness of the video clips was minimally adjusted.

We used commercially available software for quantifying facial expressions (FaceReader; version 7; Noldus Information Technology Inc). This software was developed based on a quantitative evaluation method called the Facial Action Coding System [30], which describes visually identifiable facial muscle movements as “action units,” identifies the intensity of a basic emotional state, and outputs time-series data sets comprising expression intensity values from 0 to 1 for each of the 7 facial expression elements (ie, neutral, happy, sad, angry, surprised, scared, and disgusted) on a continuous scale with high accuracy [27,31,32]. This method is advantageous in that it can quantitatively evaluate facial expressions conveniently with good reproducibility, unlike subjective evaluations, as discussed previously [33]. The software provides 5 face models (General, General61, Children, East Asians, and Elderly) that correspond to the data sets used in the algorithm training. We used the East Asian face model according to the software specifications.

Additionally, we used the calibration function provided by the software to minimize person-specific biases due to facial wrinkles or light effects. The reference manual of FaceReader explains that this function removes biases in the 7 facial expression elements but does not increase the intensity. For each participant, the resting facial image in the first part of recording was used as the calibration image for all the video clips of the participant. In cases where certain facial expression elements were detected in the neutral control image, those expression elements were corrected in the images to be analyzed. The degree of successful face recognition was evaluated for every video clip, and the session was excluded from further analysis if both images from the 2 cameras had a low proportion of successful frames (<20%), with reference to previous studies [32,34].

**Clinical Assessments**

In addition to basic characteristics, overall cognitive function was assessed using the Japanese version of the Mini-Mental State Examination (MMSE-J) [35] and the Japanese version of the Montreal Cognitive Assessment (MoCA-J) [36]. The Barthel index was used to assess performance on 10 basic activities of daily living (ADL), which tend to deteriorate in people with dementia. The total score ranges from 0 (worst; all dependent) to 100 (best; all independent). The Dementia Behavior Disturbance Scale (DBDS) [37,38] was used to assess the severity of neuropsychiatric symptoms. The scale evaluates a total of 28 items on a 5-point scale from 0 (not at all) to 4 (always) in terms of the frequency of the behavioral disturbances typically seen in persons with dementia, such as wandering, agitation, and aggression, and is scored from 0 (best) to 112 (worst). The questionnaire format allowed caregivers to answer the questions easily. Finally, to assess hearing disability, which may affect the response to auditory stimulation, 10 items from the Questionnaire on Hearing [39] were used to assess the severity of hearing loss in daily life. These 10 questions set up specific situations of hearing speech or environmental sounds in daily life and were to be answered on a 5-point scale from 1 (always able to hear) to 5 (never able to hear). The total score ranges from 10 (best) to 50 (worst).

The MMSE-J and MoCA-J were administered by skilled occupational therapists on different days. The Barthel index, DBDS, and the Questionnaire on Hearing were scored by nursing home staff members who were sufficiently familiar with the participants.

**Statistical Analyses**

The averages of the emotions expressed in response to the robot (ie, happy, sad, angry, surprised, scared, and disgusted) were compared to each other using the nonparametric Wilcoxon signed-rank test, with the P values multiplied by the number of tests according to the Bonferroni method. The correlations between subjective and objective emotional assessments were also examined using Spearman correlation coefficients. The average expression intensity values during the 30 seconds for each sensory stimulus obtained by facial expression analysis were compared for all 3 patterns (visual only; visual and auditory; or visual, auditory, and tactile) using the Friedman test with the Wilcoxon signed-rank test as a post hoc test. The average of the 3 robot presentation sessions was used in this part of the analysis. Additionally, focusing on positive emotions, the expression intensity values for “happy” were compared between the different types of robots using the Friedman test with the Wilcoxon signed-rank test as a post hoc test. Statistical analyses were performed using STATA/SE (version 13.1; StataCorp). Any P value less than .05 was considered statistically significant.

**Results**

Table 1 presents the demographic characteristics of the participants. Eleven participants (38%) had been diagnosed with Alzheimer disease by their physician, 5 (17%) had dementia with Lewy bodies, and 13 (45%) met the criteria for major neurocognitive disorder in the *DSM-5*, but the etiology was not specified. All the participants had cognitive decline when compared with the cutoff value of 26 on the MoCA-J [36].
Table. Overall participant characteristics (N=29).

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years), mean (SD; range)</td>
<td>88.7 (6.2; 71-98)</td>
</tr>
<tr>
<td>Gender, n</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>2</td>
</tr>
<tr>
<td>Female</td>
<td>27</td>
</tr>
<tr>
<td>Type of disease, n</td>
<td></td>
</tr>
<tr>
<td>Alzheimer disease</td>
<td>11</td>
</tr>
<tr>
<td>Dementia with Lewy bodies</td>
<td>5</td>
</tr>
<tr>
<td>Not specified</td>
<td>13</td>
</tr>
<tr>
<td>Years of education, mean (SD; range)</td>
<td>9.7 (2.2; 6-13)</td>
</tr>
<tr>
<td>MMSE-J, mean score (SD; range)</td>
<td>18.2 (5.1; 11-28)</td>
</tr>
<tr>
<td>MoCA-J, mean score (SD; range)</td>
<td>11.8 (4.9; 2-24)</td>
</tr>
<tr>
<td>Barthel index, mean score (SD; range)</td>
<td>66.0 (24.8; 10-95)</td>
</tr>
<tr>
<td>10 items from the Questionnaire on Hearing, mean score (SD; range)</td>
<td>26.6 (8.6; 14-50)</td>
</tr>
<tr>
<td>DBDS, mean score (SD; range)</td>
<td>13.4 (9.6; 0-40)</td>
</tr>
</tbody>
</table>

aA total of 4 patients were taking medication for dementia.
bA total of 2 patients were taking medication for dementia.
cMMSE-J: Japanese version of Mini-Mental State Examination.
eDBDS: Dementia Behavior Disturbance Scale.

The average subjective emotional assessments (5-point Likert scale) and objective expression intensity values for the expression elements across all of 3 types of robots are shown in Table 2. Facial analysis failed to detect any action units or emotional elements in 1 of the 29 participants. Also, 1 video clip was excluded from the analyses because of a low proportion of successful frames (participant 13; doll-type robot; visual, auditory, and tactile stimuli presented). After excluding these video clips, the overall percentage of the video frames analyzable by the software was 81.4%. Among the self-reported emotions, “happy” was significantly the most common (happy vs surprised: \( P=.01 \); happy vs sad, angry, scared and disgusted: \( P<.001 \)), and facial analysis–detected emotions showed the same trend in that the values of “happy” were significantly the most common among the 6 emotional elements assessed (happy vs all others: \( P<.001 \)). Additionally, the correlations between subjective and objective emotional assessments were significant for “happy,” “sad,” and “surprised,” though the correlation coefficients were interpreted as slight or low. Taking the value of “happy” as an example, as shown in Figure 3, there were cases where the objective value was detected as high even when the subjective feelings were reported as low.

Table. The relationship between subjective and objective emotional assessments.

<table>
<thead>
<tr>
<th>Subjective (5-point Likert scale; n=29)</th>
<th>Objective (expression intensity values; n=28)</th>
<th>( \rho )</th>
<th>( P ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD)</td>
<td>Range</td>
<td>Mean (SD)</td>
<td>Range</td>
</tr>
<tr>
<td>Happy</td>
<td>3.6 (1.0)</td>
<td>0.18 (0.16)</td>
<td>0-0.73</td>
</tr>
<tr>
<td>Sad</td>
<td>2.1 (0.8)</td>
<td>0.07 (0.09)</td>
<td>0-0.61</td>
</tr>
<tr>
<td>Angry</td>
<td>1.9 (0.8)</td>
<td>0.07 (0.11)</td>
<td>0-0.77</td>
</tr>
<tr>
<td>Surprised</td>
<td>3.1 (1.1)</td>
<td>0.09 (0.11)</td>
<td>0-0.62</td>
</tr>
<tr>
<td>Scared</td>
<td>1.9 (0.7)</td>
<td>0.03 (0.05)</td>
<td>0-0.36</td>
</tr>
<tr>
<td>Disgusted</td>
<td>2.0 (1.0)</td>
<td>0.05 (0.06)</td>
<td>0-0.34</td>
</tr>
<tr>
<td>Neutral</td>
<td>N/A</td>
<td>0.46 (0.13)</td>
<td>0.14-0.85</td>
</tr>
</tbody>
</table>

aN/A: not applicable.
Figure 3. Comparison of the objective expression intensity for “happy” at each grade of subjective emotional assessment (5-point Likert scale). Error bars indicate SDs, the lines within the boxes indicate medians, and the dots indicate outliers that were more than a quarter of the range × 1.5 away from the first or third quartile.

Figure 4 shows a comparison of 3 different patterns of sensory stimuli among all participants. There were significant differences in the values for “neutral” (P<.001), “happy” (P<.001), “sad” (P=.001), “scared” (P=.04), and “disgusted” (P<.001) among the 3 patterns. Post hoc analyses revealed that the values for “happy” significantly increased in the pattern with visual, auditory, and tactile stimuli (median score 0.21, IQR 0.09-0.35) compared to the patterns with visual stimulus only (median score 0.10, IQR 0.03-0.22; P<.001) and with both visual and auditory stimuli (median score 0.10, IQR 0.04-0.23; P<.001).

The values for “sad” (with visual, auditory, and tactile stimuli: median score 0.05, IQR 0.01-0.11) and “disgusted” (with visual, auditory, and tactile stimuli: median score 0.04, IQR 0.02-0.10) exhibited the same trend, though both of these values were significantly smaller than those for “happy” (P<.001). In contrast, the values for “neutral” and “scared” significantly decreased in the pattern with visual, auditory, and tactile stimuli compared with the other 2 patterns. However, the intensity of each emotional element did not change linearly over time.
Figure 4. Comparison of all the expression intensities under the 3 different patterns of sensory stimulation. Error bars indicate SDs, the lines within the boxes indicate medians, and the dots indicate outliers that are more than a quarter range × 1.5 away from the first or third quartile. The values in the upper right of each group represent the overall \( P \) values in the Friedman test. The values above the boxplots for the nonresponder group represent the \( P \) values in the post hoc test (the Wilcoxon signed-rank test). \( P \) values less than .05 are denoted in bold. V: visual stimulus only; V+A: visual and auditory stimuli; V+A+T: visual, auditory, and tactile stimuli.

When comparing the expression intensity values for “happy” between the different types of robots for the same participants and the same sensory stimuli, no statistical differences were found, as depicted in Figure 5A. When comparing the expression intensity values for “happy” between the different sensory stimuli for the same robot type, no statistically significant differences were found for robot A (humanoid-type; \( P=.48 \)), while robot B (doll-type; \( P<.001 \)) and robot C (animal-type; \( P=.03 \)) had large, significant values in the pattern with visual, auditory, and tactile stimuli compared with the other 2 patterns, as shown in Figure 5B.
Figure 5. Comparison of the expression intensity for “happy” between the different types of robots. (A) Comparison between 3 types of robots with the same pattern of sensory stimulation. (B) Comparison within each robot type using the 3 different patterns of sensory stimulation. Error bars indicate SDs, the lines within the boxes indicate medians, and the dots indicate outliers that are more than a quarter of the range × 1.5 away from the first or third quartile. The values in the upper right of each group represent the overall P values from the Friedman test. The values above the boxplots for robots B and C represent the P values from the post hoc test (the Wilcoxon signed-rank test). P values less than .05 are denoted in bold.

Discussion

Principal Results

This study aimed to quantitatively evaluate the psychological and emotional responses of people with dementia to complex sensory stimuli provided by socially assistive robots. To summarize, the values for “happy” were found to be the largest in the overall response to the robots among the emotional elements in both the self-recorded assessment and objective facial expression analysis. However, correlations between the subjective and objective emotional assessments were found to be relatively low or insignificant. When comparing the 3 different patterns of sensory stimulation, there was a significant increase in some of the expression intensity values when all stimuli (visual, auditory, and tactile) were presented. In the comparison between different types of robots, focusing on “happy,” there was a significant increase when all stimuli were presented by a doll-type robot and by an animal-type robot but not by a humanoid-type robot.
The emotional responses revealed by the objective facial expression analysis showed the same trend as the participants’ subjective assessment, with happy emotions having the largest values in both assessments. This result suggests that the facial expression analysis system can successfully detect the expressions that emerge in people with dementia, which is consistent with a previous study [40].

However, considering that the facial expression intensity was detected as relatively high in some cases where the subjective rating was moderate, the facial expression analysis may be able to capture emotions that are not sufficiently represented by a subjective 5-point Likert scale. In fact, a previous study supports the appropriateness of facial expression analysis as a nonverbal pain assessment for people with dementia when they lack the ability to self-report [41]. Therefore, facial expression analysis may also be useful as an accurate nonverbal assessment of emotions for people with dementia.

Although some studies have already demonstrated positive responses in people with dementia when presented with socially assistive robots [42,43], this study is the first to objectively and quantitatively describe emotional responses using facial expression analysis. Additionally, these results are inconsistent with those of Goda et al [19], who found that a 5-minute talk session with a socially assistive robot caused stress in people with dementia. In contrast to this study, the socially assistive robot’s interaction in the study by Goda et al [19] was mainly through verbal communication; thus, it was inferred that these stimuli were burdensome for people with dementia, who tend to have a decline in verbal communication skills. The positive expressions shown in this study are considered to have been observed as responses to sensory stimuli and not verbal stimuli.

Regarding the relationship between the type of sensory stimulus and emotional responses, the expression intensity values for “happy,” “sad,” and “disgusted” significantly increased with increasing sensory stimuli. Particularly, happy emotions were revealed to be elicited most strongly in persons with dementia when different varieties of sensory stimuli were presented. With respect to the values for “sad” and “disgusted,” we consider it natural that participants became somewhat careful or timid toward unknown experiences when interacting closely with robots. However, these emotional elements in automated analysis should be interpreted with caution, as they can occasionally include other negative emotions, such as fear [44], especially when accompanied by facial movements around the lips or jaws [31]. Nevertheless, the values for these negative emotions remain small compared to the value for “happy,” showing that positive emotions were dominant when visual, auditory, and tactile sensory stimuli were offered.

Notably, this increase in positive emotional responses may include changes over time, because the sensory stimuli were added over time in this study. Given the procedures in this study, the results could have been influenced by the novelty effect [20], which is derived from curiosity toward a new experience. However, the intensity of each emotional element did not change linearly when sensory stimuli were added over time. Moreover, the addition of an auditory stimulus did not significantly increase the expression intensity value, whereas the addition of a tactile stimulus did. These findings suggest the importance of adding tactile stimuli to visual and auditory stimuli. This is plausible considering that tactile information connects through several pathways to the insular cortex, which evokes emotional responses [45,46]. The finding that tactile stimuli evoked positive feelings is supported by previous studies that showed that haptic or tangible input was effective in helping people with dementia understand and adapt to their surroundings [47,48]. Regarding auditory stimuli, the results of the Questionnaire on Hearing did not indicate that the participants had very good hearing, suggesting that the effects of auditory stimuli on emotion elicitation may have been relatively small. Since hearing loss is associated with the risk of developing dementia [49], and a high percentage of people with dementia actually have hearing loss [50,51], this issue may also be of great clinical importance.

Furthermore, it is noteworthy that a significant increase in expression intensity values for “happy” with increased sensory stimuli was observed for doll-type robots and animal-type robots, but not for humanoid-type robots. One of the distinctive characteristics of the doll-type and animal robots used in this study was that they were covered by soft, fur-like materials. The importance of soft materials is commonly discussed in the field of soft robotics for medical use or human assistance [52,53]. Softness is considered effective not only in terms of safety for the human body but also in terms of the imitation of reality or the creation of familiarity [54] and emotional processing [55]. However, the emotional effects of various tactile sensations in people with dementia have not yet been studied; consequently, given the findings of this study, the effectiveness of soft tactile stimuli in dementia care may be worth exploring in future research.

Limitations
This study has a few limitations. First, it included a small number of participants and a single experimental group. Changes in facial expressions were reliably detected by using the participants’ resting states as controls. However, another study design, such as a randomized controlled trial, is required to confirm these effects more clearly. Second, the generalizability of our findings may be limited, as most of the participants in this study were female. Previous studies report that there are gender differences in emotional responses to some types of sounds [56], emotion expression [57], and emotion regulation [58], although gender and facial expression have been reported to have no significant correlation [40]. Further research with male participants will be needed to reveal possible gender differences in responses toward social robots. Third, since there is no prior literature that has identified a minimum detectable change or a minimal clinically important difference for expression intensity, the clinical significance of the changes in expression intensity values demonstrated in this study needs to be explored further. Finally, this study only investigated the immediate responses to socially assistive robots, with patients allowed to interact with each robot for only 1.5 minutes in total. However, in real clinical settings, people with dementia might express more diverse patterns of responses, using these robots as they would like. Moreover, any enthusiasm resulting from the novelty effect may diminish over time. Thus, further
investigation is required to reveal the long-term emotional effects of socially assistive robots on people with dementia, including variability in positive responses over several hours or days of use and the effects of these robots on their neuropsychiatric symptoms.

**Conclusions**

This study quantitatively examined the emotional reactions of people with dementia to socially assistive robots. The expression intensity values, especially the values for “happy,” significantly increased with multiple sensory stimuli, including visual, auditory, and tactile stimuli. Therefore, this study shows that socially assistive robots may be more effective in arousing positive emotions when multiple sensory stimuli are involved. Further studies, including randomized controlled trials, are required to further explore the effectiveness of and the optimal methods for using socially assistive robots in dementia care.

Acknowledgments

EO, AO, KK, KM, and SH conceptualized and designed the study. EO collected and analyzed data and drafted the manuscript. MK participated in data collection, data interpretation, and manuscript editing. AO, KK, YO, and SU participated in data interpretation and manuscript editing. IK participated in data interpretation and critical revisions of the manuscript. All authors have read and approved the final manuscript. This study was supported by the Japan Health Research Promotion Bureau Research Fund for Young Investigators (J1H2021-Y-11). We sincerely thank the staff of Wakyokai Nagomi-no-Sato and Jinshikai Ruminasu Obu (nursing homes for older people) and the Wellness Valley Promotion Office (Obu City) for their cooperation. We also thank Naho Hashimoto, Ai Sugiyama, Michiko Chiso, and Ayumi Ogura for their technical support.

Conflicts of Interest

None declared.

References


31. Smedegaard CV. Why a proper investigation of novelty effects within SHRI should begin by addressing the scientific plurality of the field. Front Robot AI 2022 May;9:741478. [doi: 10.3389/frobt.2022.741478] [Medline: 35719207]


Abbreviations

ADL: activities of daily living
CDR: Clinical Dementia Rating
DBDS: Dementia Behavior Disturbance Scale
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
MMSE-J: Japanese version of Mini-Mental State Examination
MoCA-J: Japanese version of Montreal Cognitive Assessment

https://aging.jmir.org/2024/1/e52443

JMIR Aging 2024 | vol. 7 | e52443 | p.872
(page number not for citation purposes)
Promoting Personalized Reminiscence Among Cognitively Intact Older Adults Through an AI-Driven Interactive Multimodal Photo Album: Development and Usability Study

Xin Wang, PhD; Juan Li, PhD; Tianyi Liang, BSc; Wordhul Hasan, BSc; Kimia Tuz Zaman, BSc, MSc; Yang Du, PhD; Bo Xie, PhD; Cui Tao, PhD

1Department of Computer Science, North Dakota State University, Fargo, ND, United States
2Department of Computer Systems and Software Engineering, Valley City State University, Valley City, ND, United States
3School of Nursing, The University of Texas at Austin, Austin, TX, United States
4School of Information, The University of Texas at Austin, Austin, TX, United States
5D Bradley McWilliams School of Biomedical Informatics, UTHealth Houston, Houston, TX, United States

Corresponding Author:
Juan Li, PhD
Department of Computer Science
North Dakota State University
Quentin Burdick Building Room 258
1320 Albrecht Boulevard
Fargo, ND, 58105
United States
Phone: 1 7012318562
Fax: 1 7012318255
Email: J.Li@ndsu.edu

Abstract

Background: Reminiscence, a therapy that uses stimulating materials such as old photos and videos to stimulate long-term memory, can improve the emotional well-being and life satisfaction of older adults, including those who are cognitively intact. However, providing personalized reminiscence therapy can be challenging for caregivers and family members.

Objective: This study aimed to achieve three objectives: (1) design and develop the GoodTimes app, an interactive multimodal photo album that uses artificial intelligence (AI) to engage users in personalized conversations and storytelling about their pictures, encompassing family, friends, and special moments; (2) examine the app’s functionalities in various scenarios using use-case studies and assess the app’s usability and user experience through the user study; and (3) investigate the app’s potential as a supplementary tool for reminiscence therapy among cognitively intact older adults, aiming to enhance their psychological well-being by facilitating the recollection of past experiences.

Methods: We used state-of-the-art AI technologies, including image recognition, natural language processing, knowledge graph, logic, and machine learning, to develop GoodTimes. First, we constructed a comprehensive knowledge graph that models the information required for effective communication, including photos, people, locations, time, and stories related to the photos. Next, we developed a voice assistant that interacts with users by leveraging the knowledge graph and machine learning techniques. Then, we created various use cases to examine the functions of the system in different scenarios. Finally, to evaluate GoodTimes’ usability, we conducted a study with older adults (N=13; age range 58-84, mean 65.8 years). The study period started from January to March 2023.

Results: The use-case tests demonstrated the performance of GoodTimes in handling a variety of scenarios, highlighting its versatility and adaptability. For the user study, the feedback from our participants was highly positive, with 92% (12/13) reporting a positive experience conversing with GoodTimes. All participants mentioned that the app invoked pleasant memories and aided in recollecting loved ones, resulting in a sense of happiness for the majority (11/13, 85%). Additionally, a significant majority found GoodTimes to be helpful (11/13, 85%) and user-friendly (12/13, 92%). Most participants (9/13, 69%) expressed a desire to use the app frequently, although some (4/13, 31%) indicated a need for technical support to navigate the system effectively.
Conclusions: Our AI-based interactive photo album, GoodTimes, was able to engage users in browsing their photos and conversing about them. Preliminary evidence supports GoodTimes’ usability and benefits cognitively intact older adults. Future work is needed to explore its potential positive effects among older adults with cognitive impairment.

(JMIR Aging 2024;7:e49415) doi:10.2196/49415

KEYWORDS
aging; knowledge graph; machine learning; reminiscence; voice assistant

Introduction

As the proportion of older individuals rapidly grows, an increasing number of older individuals are becoming concerned about their physical and mental well-being [1]. Steptoe et al [2] found that a decline in health is associated with a negative psychological state. In addition, aging is closely associated with various psychosocial stress factors, such as loneliness, personal losses, and lower socioeconomic status [3]. These factors may increase the risk of developing mental health disorders. In recent years, increasing evidence suggests that psychological well-being could be a potential asset for healthy aging [4].

Reminiscence therapy stands as a profound approach, rooted in a deep understanding of cognitive and emotional processes, designed to elevate the psychological well-being of older adults. This therapy transcends the mere recall of life histories, encompassing both oral and written narratives, in a multisensory journey that engages sight, sound, taste, touch, and smell [5-7]. It extends beyond a mere collection of activities, encompassing the contemplation of photographs, immersion in music, and the sharing of narratives about pivotal life events [5-8]. The underlying success of reminiscence therapy lies in its capacity to stimulate long-term memory, playing a pivotal role in fostering overall well-being, an attribute particularly invaluable for those grappling with short-term memory challenges [9]. As individuals review and discuss evocative materials, they embark on a path to not only retrieve memories but also enhance their self-esteem, nurture interpersonal skills, and enrich their psychosocial well-being [4,10]. This is rooted in the therapy’s ability to harness the profound impact of these sensory stimuli. Significantly, the American Psychological Association recognizes the therapeutic potential of reminiscence therapy, attesting to its ability to ameliorate mental health conditions, elevate mood, and mitigate agitation, especially among individuals coping with Alzheimer disease or dementia [5].

Research by Tam et al [4] indicates that reminiscence intervention not only produces positive effects among older adults with dementia but also benefits cognitively intact older adults. For instance, it reduces the depressive symptoms of cognitively intact older adults, significantly improves their life satisfaction, and promotes their self-esteem, psychological well-being, and happiness.

Reminiscence interventions, whether administered within health care facilities such as hospitals, assisted living communities, or nursing homes, or within the familial cocoon of private homes, are bolstered by the skillful orchestration of trained professionals and caregivers [5,7,11]. In health care settings, psychologists, social workers, and specialists in geriatric care often take the helm in conducting these interventions, drawing upon a wealth of materials that hold personal significance for older adults. The tailored application of these materials, including photos, videos, and cherished objects, becomes a key facet in rekindling memory and reminiscence. These interventions can be adjusted to individual or group settings, finely attuned to the unique needs and preferences of older adults [5,11,12]. In private residences, reminiscence interventions are executed with dedication by family members, caregivers, or volunteers [5]. They use similar materials as those found within health care facilities, including familial photographs and cherished keepsakes, as a conduit to memory stimulation and the initiation of reminiscing. Nevertheless, conducting these interventions at home can be challenging due to a shortage of human resources [11]. Caregivers and family members, juggling numerous responsibilities, may find their time and resources stretched thin, underscoring the need for innovative approaches, such as the one proposed in this study, to bridge this gap.

To overcome the challenges of limited resources and specialized training in performing reminiscence interventions, flexible and effective computer-based interventions are highly beneficial [13-15]. We designed, developed, and tested GoodTimes, a personalized interactive multimodal photo album mobile app for cognitively intact older adults. It can be used on smartphones and tablets, providing older adults with on-the-go access to their photos as well as those provided by family members. This intelligent digital photo album allows users to browse and query photos in various orders, including chronological order, by location, by event, or by specific family members. To enhance the reminiscing experience, a voice assistant (VA) interacts with the user, providing information about the picture, such as the family members in the picture, the special moment, and the location. The VA also asks questions, answers user questions, and responds to user comments, creating a setting where special memories can be remembered and enjoyed. To ensure the user’s comfort and avoid any distress, the VA will not mention anything about depression or other challenges that older adults may be facing. Older adults can use the app independently, but it is also beneficial for families, caregivers, and assisted living staff to share the mobile app with the older adult they care for. Overall, GoodTimes is a convenient and effective tool that can be used anytime, anywhere, with or without caregiver support, for cognitively intact older adults. This user study indicates that after using this app, users reported engaging in enjoyable conversations with the app.
Methods

Ethics Approval
This study was reviewed and approved by the institutional review board of NDSU. The IRB Protocol number is IRB0004419.

Overview
GoodTimes is built on artificial intelligence (AI) technologies, including image recognition, natural language processing, knowledge graph, logic, and machine learning, to provide an interactive and personalized experience for older adults. The system architecture is illustrated in Figure 1. The app can be accessed through smartphones and tablets, and users can interact with it using their voice or fingers. The VA, which uses automatic speech recognition and natural language understanding technologies, converts the user’s voice into a text request. The conversation management module then processes the user’s request through 2 steps: user intent identification and dialogue management (DM). First, the user intent identification matches the user’s text request with predefined intents and dialogue states to create an input frame. Then, the DM module executes the dialogue policy based on the dialogue state graph and updates the dialogue state.

Figure 1. System Architecture.

There are 2 types of communication requests: personal requests, which involve questions and answers related to the user’s personal information, and general requests, which do not necessarily involve personal information. Examples of personal questions include “Who is in the photo?” and “What is the story behind this photo?” General requests include knowledge not necessarily related to the user’s personal information, for example, a commonplace or a piece of widely known artwork. For instance, the AI may provide an answer about art collections in the Louvre Museum in Paris.

To handle general requests, we fine-tuned a Generative Pre-Trained Transformer (GPT; OpenAI) [16], a powerful neural language model. For personal requests, we have designed our own personal knowledge engine to provide tailored responses. Backend services of our personal knowledge engine are requested based on the user’s intent, and these services are supported by a semantics-based query engine and reasoning engine. These engines work over a knowledge graph, which is the brain of the system. The knowledge graph contains facts, relationships, and rules about photos, people, places, time, and stories. The search and reasoning engines link the dialogue with a specific photo, user profile, and context to enable personalized services. Finally, the DM module generates responses using the speech act and content selected based on the input frame.

The system maintains a library of photos that can be uploaded by the older adult user’s family members and caregivers. Metadata of a photo, such as people, animals, location, time, and special events, are also saved in the knowledge. Family members get involved in uploading, sharing, and explaining the photos, promoting collaboration, and having fun. Involving family members in the photo album can improve their relationship, help family members learn more about older adults, and facilitate reminiscence interventions for older adults, with or without caregiver support.

Constructing an Open Personal Knowledge Graph
The “brain” of the system is a comprehensive knowledge graph [17] that contains knowledge about the user and the photos. All the knowledge is represented as a graph in which data is modeled as nodes (vertices) and links (edges) between them. Nodes in our knowledge graph are normally a person, place, location, or thing, and links are how they are connected or related. Figure 2 shows an example knowledge graph about an older adult, Bob. In this graph, Bob, Alice, Cat, Person, Attraction, and Golden Gate Bridge are nodes. They are connected by many different relationships. For example, Alice is Bob’s wife, and Bob is the father of Cat. Alice, Bob, and Cat are persons. Bob visited the Golden Gate Bridge, which is an attraction. This example graph shows many things and relationships about Bob. It is used by the system to explain the story of the picture, including the people inside the picture, their relationships, where they visited, etc.
Graphs are used to model knowledge because they can encode real-world knowledge as “things” (not “strings”) and their interrelationships. This enables the system to communicate with users by analyzing what the words in a sentence actually mean rather than simply analyzing strings of characters. In addition, structuring knowledge in the form of graphs allows knowledge to evolve over time. New “things,” relationships, and external knowledge can be easily added or linked to the existing graph. In our system, we extend the core knowledge graph with external knowledge graphs, such as Wikidata [18,19], to extend our knowledge. Furthermore, reasoning and navigation can be performed over knowledge graphs.

**Knowledge Graph Construction**

First, we built a high-level ontology working as the schema of the knowledge graphs. Then, we create a knowledge graph by instantiating the ontology with detailed instances retrieved from uploaded photos (with metadata) and user surveys and external knowledge graphs, such as Wikidata. Using ontology would allow logical inference for retrieving implicit knowledge rather than only allowing queries requesting explicit knowledge. We proposed a “Who-What-When-Where” model as the foundation of this album ontology. “Who” represents the person in or not in the photo but related. “What” points to the story related to a photo. “When” specifies the time when the photo was taken, which can be a date, a social occasion, or a historical monument. “Where” describes the place where the photo is taken. Figure 3 shows a major part of the ontology. This ontology is instantiated with instances through photo metadata extraction, tagging extraction, image recognition, and social media extraction.

The ontology serves as schema-level knowledge used to instantiate instances or individuals, thereby creating a comprehensive knowledge graph. Instance information is collected through various means, such as automatically extracting metadata from photos or through manual input or voice-based question and answer. Through these processes, a detailed knowledge graph like the one shown in Figure 2 can be generated.
Semantic Rule Definition

In order to facilitate logical inference and gain new insights from the knowledge graph, we have established a set of rules and used logical reasoning over the ontology. We provide a few examples of these rules below, noting that some are specified within the ontology itself, while others are created separately using Semantic Web Rule Language [20-22]. For clarity, we present all rules using the same first-order logic [22] format.

Relatives:

If x is the husband of y, then y is the wife of x, and vice versa:

$$\forall x, y \text{ Husband}(x, y) \leftrightarrow \text{Wife}(y, x)$$

If x is the mother or father of y, then x is also the parent of y, and vice versa:

$$\forall x, y (\text{Mother}(x, y) \lor \text{Father}(x, y)) \leftrightarrow \text{Parent}(x, y)$$

If z is the parent of both x and y, and x is not the same as y, then x and y are siblings:

$$\forall x, y, z (\text{Parent}(z, x) \land \text{Parent}(z, y) \land x \neq y) \Rightarrow \text{Sibling}(x, y)$$

If x is the parent of y, and y is the parent of z, then x is the grandparent of z:

$$\forall x, y, z (\text{Parent}(x, y) \land \text{Parent}(y, z)) \Rightarrow \text{Grandparent}(x, z)$$

Social Relations:

Time:

$$\forall x, y, z \text{ Time}(x) \land \text{Time}(y) \land \text{Time}(z) \land \text{Before}(x, y) \land \text{Before}(y, z) \Rightarrow \text{Before}(x, z)$$

$$\forall x, y \text{ Time}(x) \land \text{Time}(y) \land \text{Before}(x, y) \Rightarrow \text{After}(y, x)$$

Location:
... Photo co-occurrence Relations:

\[
\forall p, l, t, ph \ (\text{Person}(p) \land \text{Location}(l) \land \text{Time}(t) \land \text{Photo}(ph) \land \text{PersonInPhoto}(p, ph) \land \text{PhotoTakenAtLocation}(ph, l) \land \text{PhotoTakenTime}(ph, t)) \Rightarrow \text{Visited}(p, l, t, ph)
\]

\[
\forall p_1, p_2, l, t, ph \ (\text{Person}(p_1) \land \text{Person}(p_2) \land \text{Visited}(p_1, l, t, ph) \land \text{Visited}(p_2, l, t, ph)) \Rightarrow \text{VisitedTogether}(p_1, p_2, l, t)
\]

Dialogue Management

The photo album’s VA is a machine learning-based system that enables users to engage with it through natural conversation. The VA is capable of understanding user intents from free text, answering questions, and asking questions for a specific purpose. To promote reminiscence intervention, GoodTimes guides users in recalling their memories by asking them photo-related questions and responding based on their answers. The VA facilitates personalized questions during the dialogue flow and can route natural conversations with users. A photo album with a VA can help older adults feel more connected to their past and present, providing them with a sense of familiarity and comfort, which is especially important for older adult users. Our dialogue flow management includes the following key points to facilitate engaging interactive conversation: VA-driven conversation, intent recognition, context management, personalization, and empathy incorporation.

**VA-Driven Conversation**

In our app, the VA initiates and guides the conversation. The conversation between the VA and the user begins with a friendly greeting or prompt, followed by a series of questions that revolve around the “Who-What-When-Where” themes but are not limited to them. These questions are designed to elicit specific information about the photos from the user and jog their memory. We use techniques including contextual prompts, confirmation prompts, and error handling to let the VA control the dialogue flow.

The VA provides prompts or suggestions to the user based on the current context of the conversation. For example, if the user talks about a specific photo, the VA can suggest related topics or questions to keep the conversation flowing smoothly. The VA uses confirmation prompts to confirm the user’s intent or response to a question. This is useful when the VA needs to verify information before moving on to the next question or action. When the user provides incorrect or invalid input, the VA will provide appropriate responses, including rephrasing a question or prompt, asking for clarification, or providing an explanation of what the VA is looking for. By using these techniques, the VA can guide the conversation in a way that ensures the user provides the necessary information while keeping the conversation under control.

**Intent Recognition**

There are 2 ways to identify the user’s intent. One is to fine-tune a GPT to let it specify the intent or use our designed intent identification model (IIM). In our implementation, we used our own IIM as the main method, as GPT is more expensive. In our IIM, the VA uses the natural language processing algorithm, part of speech tagging, to break down a sentence or phrase into its constituent parts, such as nouns, verbs, and adjectives. Then, it uses named entity recognition to extract important information such as the user’s intent, entities (relevant keywords or phrases), and context from these components. We use machine learning algorithms (eg, our previous proposed algorithm) to analyze the user’s input and match it with the most relevant intent. To train the model, we provide sample user inputs and assign them to specific intents. The VA then uses these examples to learn patterns in the data and improve its ability to recognize user intent over time.

**Context Management**

The VA also keeps track of the conversation’s context, including previous statements made by the user and the VA’s responses. This helps to ensure that the VA’s responses are relevant to the current conversation. Context management in the VA of the interactive photo album is critical to providing a seamless and personalized user experience. VA uses a context stack to manage the context of the conversation. For example, suppose the user is looking at a photo of a trip to Paris taken in front of the Louvre Museum. In that case, the VA can use this information to provide related suggestions or ask follow-up questions, such as “Did you see Leonardo da Vinci’s Mona Lisa in Louvre?” These questions are generated by prompting GPT using our knowledge graph and previous conversation history as context. The VA also needs to be able to handle changes in context, such as if the user switches to talking about a different topic. In such cases, the VA must recognize the change in context and adjust its responses accordingly. Overall, effective context management is crucial to creating a personalized and engaging experience for users interacting with the VA in the interactive photo album.

**Personalization**

The VA personalizes the conversation by considering multiple factors, including the user’s preferences, personal profile such as name, age preferences, and conversation history. The very basic form of personalization is addressing the user by name to make the conversation more personal and engaging. In addition, the VA will use knowledge in the knowledge graph to address people or things in the photo. For example, the VA will use the information stored in the knowledge graph to refer to people or things depicted in the photos. For instance, if the user’s mother is shown in the photo, the VA may address her as “your mother, Susan,” as her name is known from the knowledge graph. Similarly, if the user’s pet dog is in the photo, the VA may refer to the dog by its name, “Buddy.” Additionally, suppose the knowledge graph indicates that the user has a close relationship with a particular person. In that case, the VA can refer to them with a personal term, such as “your dear friend, John.” This personalization can enhance the user’s experience and create a more natural and engaging conversation. Also, the VA uses the conversation history to tailor the conversation.

https://aging.jmir.org/2024/1/e49415

JMIR Aging 2024 | vol. 7 | e49415 | p.879

JMIR Aging 2024 | vol. 7 | e49415 | p.879

(page number not for citation purposes)
the user has previously shown a preference for a particular type of photo or event, the VA can use this information to recommend similar photos or events.

**Empathy Incorporation**

The VA incorporates empathy into dialogue to create a more natural and engaging conversation. This involves understanding and responding to the user’s emotional state, using appropriate tone and language, and showing concern for the user’s needs and feelings. Older adults and people with Alzheimer disease may have difficulty understanding complex sentences or abstract concepts [26]. Our VA uses simple, clear language to make sure they understand what the VA is saying. The VA always tries to be patient and understanding when asking older adults about photos. The VA gives positive feedback when the user answers questions correctly or remembers important information. If they provide incorrect answers, the VA will gently correct them and provide additional context or information. It is also important to repeat questions if they are not answered correctly, as older adults may need more time to process and remember information. Asking related questions, such as about memories of a trip shown in the photo or the hobbies of a person in the photo, can also be helpful in stimulating memories and encouraging conversation. Overall, we try to create a comfortable and positive environment for older adults to share their memories and stories.

**Conversation Using Knowledge Graph**

The knowledge graph is the source of information for conversing with users and is stored in Neo4j [27-29], a graph database that the VA uses to ask and answer questions about photos. Cypher [27-29], Neo4j’s query language, is used by the VA to navigate the graph and generate questions and responses. Natural language queries and answers from users are converted into Cypher queries. For example, if a user asks, “Who is in this photo?” the VA can use this to create a Cypher query that retrieves all people in the photo. To generate photo-related questions, Cypher first locates a specific photo node based on certain criteria. Relevant properties are then extracted from this node to generate a question, with the property value serving as the standard answer. Cypher can also query for a relationship linked to this node, creating a directional triple consisting of the photo node, a relationship, and another node. The knowledge graph uses directional triples (node-relationship-node) to express semantic data as subject-predicate-object. For more complex questions with multiple bindings, Cypher enables searching for nearby nodes and relationships by limiting the number of hops and specific relationships.

To keep the conversation engaging, the VA can ask follow-up questions related to the previous topic. For instance, if the previous query was about Mary, the VA might ask, “Do you remember Mary’s hobby?” by using the previous query result (ie, Mary) as the subject or object of a new triple. The new question is generated based on this triple, with the other node or relationship serving as the question and the remaining element in the triple as the standard answer. In addition, GPT [30,31] is used to generate related questions and content, which will be discussed further later on.

**GPT as a Complement**

GPT is the state-of-the-art large language model used for various natural language processing tasks, including “question answering.” One of the main advantages of GPT is its large amount of pretrained knowledge, which enables it to understand and generate natural language text with high accuracy. However, the cost of using GPT could be prohibitive for some applications, including our system. Despite this limitation, we still benefited from GPT’s capabilities by using it to provide complementary functions. For example, we used GPT to help identify user intent if our IIM was not confident about its result. Moreover, GPT was used to enhance the conversational experience by generating follow-up questions and responses. For instance, if a user asks a question about a particular topic, GPT can generate related questions or statements that may help the user explore the topic further.

To ensure that the responses generated by GPT are relevant and accurate, we provided appropriate context for the conversation. One way to achieve this is by leveraging our local knowledge graph and conversation history as a source of relevant information. By converting the knowledge graph database into a triple format (subject, predicate, and object), we could index the triple file using semantic embeddings, which represent text data in a continuous vector space. This allows for efficient comparison and retrieval of similar text items while preserving semantic relationships between words and phrases. To generate embeddings, we used the pre-trained embedding model “Bidirectional Encoder Representations from Transformers” [32], resulting in a single vector embedding. These embeddings were then used in semantic search, enabling efficient and cost-effective searching. When a user query was received, we converted it into semantic embeddings and matched them with the embeddings of the knowledge graph using cosine similarity. This helped us identify the most relevant or related information sources in our knowledge graph that are related to the user’s query. Once we had identified the most similar embeddings in our local knowledge base, we used them as context or query input for the GPT model. This context provides GPT with relevant information from our local knowledge base, allowing it to generate relevant responses. In summary, while GPT may not be used extensively in our system due to its cost, it can still provide valuable complementary functions that enhance the user experience and help us better understand the user’s intent.

**Prototype System**

We have implemented GoodTimes as an Android-based app using a mobile app development framework called Flutter [33]. The app includes a VA that uses Google’s Dialogflow [34] to understand natural language. When the VA identifies what the user wants, it uses Google Cloud Functions to interact with a backend server built with Spring Boot [35]. This server connects to a Neo4j database that stores information in the form of a knowledge graph. We used LangChain (Harrison Chase) [36], a Python package, to integrate OpenAI’s GPT language models with the knowledge graph. Specifically, we used the GPT-3.5 Turbo model and defined the GPT language model using the LLMPredictor class and the input prompt format using the
PromptHelper class. The user has the option to interact with the app using either voice commands or touch input.

**Use-Case Study**

We conducted a use-case study to evaluate the functionality of our interactive photo album app before deploying it to real users. This involved identifying and analyzing specific scenarios in which the app could be used, such as browsing photos, sharing them with friends and family, and asking questions related or unrelated to the photo. By examining how the app was used in these scenarios, we were able to identify areas for improvement to make it more user-friendly. The creation of fictional users for the use cases involved careful consideration of relevant characteristics and demographics aligned with this study’s objectives. Factors taken into account included age, gender, cultural background, language proficiency, cognitive and physical abilities, technological familiarity, and more. This iterative process allowed for refinement based on the specific goals and requirements of our research or design project.

Let us consider a fictional character, Bob, who is 79 years old and has memory challenges that affect his ability to remember things. As part of our use case, we presented Bob with a collection of photos encompassing significant moments such as his granddaughter Sara’s birthday party, his memorable trip to Paris with his wife, and heartwarming memories with his beloved pet dog, Lucy.

**Figure 4** exemplifies the interaction between Bob and the VA, showcasing their communication regarding photos in the album. This particular example demonstrates the VA’s capability to engage with Bob and help him remember significant moments. In the conversation, Bob and the VA discuss a photo from Sara’s birthday party. Initially, Bob is unsure of the people in the photo. The VA kindly reminds him that it is Sara in the picture, which prompts Bob’s memory. The VA provides additional information about the event, refreshing Bob’s memory and filling in the gaps. The VA also helps Bob recall his trip to Paris. This interaction highlights the VA’s ability to serve as a helpful reminder and facilitator of memory recall. By engaging in conversations and providing contextual information, the VA helps Bob reminisce about important occasions and people in his life. This feature demonstrates the valuable role the VA plays in aiding individuals with memory challenges by providing gentle reminders and support.

In addition to the previously mentioned use cases, we conducted evaluations on 12 additional scenarios for our app. The diverse range of use cases demonstrated the app’s potential as a valuable tool for helping older adults cherish their memories and enhance their overall well-being.
**User Study**

We conducted a user study to evaluate the usability and features of the GoodTimes app. Participants were recruited through convenience sampling using various digital communication channels, such as phone calls, social media, and email. To accommodate participants’ preferences, the study sessions were conducted in person at locations including their homes, local coffee shops, or parks. These sessions were conducted individually, allowing for personalized interaction and feedback.

At the beginning of each session, participants were provided with an overview of the study objectives and the app’s key features. Written consent was obtained from each participant, following the approved study protocol by the institutional review board of North Dakota State University. A video demonstration was presented, and participants were given approximately 30 minutes of hands-on practice with the app. A researcher was present to address any questions during the session. Following the interaction, participants completed a set of questionnaires to provide feedback on their experience, which took approximately 15 minutes to complete. This approach enabled us to gather comprehensive insights into the app’s usability and user experience.

The questionnaire consists of 2 main parts. The first part focused on participants’ experiences with the app’s features, including their interactions with the intelligent agent and their emotional responses to the digital photo gallery. They rated their experiences with 5 statements on a Likert scale ranging from “strongly agree” to “strongly disagree.” The survey included questions developed by the research team based on previous research and expert opinions.

The second part focused on the overall usability of the app, and open-ended questions were used to identify areas for improvement. In this part, we used a subset of the System Usability Scale (SUS) to assess the app’s usability. The SUS is a widely recognized and validated instrument for evaluating the usability of interactive systems. We selected 5 relevant SUS questions out of a total of 10 and included them in our survey.

In our evaluation of the app’s usability and user experience, we used statistical analyses to objectively understand participant feedback. Specifically, we used a 1-sample 2-tailed t test, a parametric test used to determine if the sample mean significantly differs from a known or hypothesized population mean. Given that our data was interval in nature and derived from Likert-scale responses, the 1-sample t test was particularly apt. The hypothesized population mean in our context was the neutral point on our scale, allowing us to discern whether participants’ responses significantly leaned toward agreement or disagreement.

**Results**

**Participants**

The user study included a total of 15 participants, with 13 of them completing the survey. Out of the 12 studies conducted, 11 were in-person sessions, while 1 was conducted on the web through Zoom (Zoom Video Communications) as per the participant’s preference. The participant demographics are summarized in Table 1. Among the participants, 8 fell in the age range of 55-64 years, 2 were in the age range of 65-74 years, and 3 were in the age range of 75-84 years. In terms of gender distribution, there were 8 male participants and 5 female participants. Regarding education, 3 participants held a PhD degree, 4 participants had a master’s degree, 3 participants had a bachelor’s degree, 1 participant had a college degree, and 2 participants had completed high school.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Participants, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
</tr>
<tr>
<td>55-64</td>
<td>8 (62)</td>
</tr>
<tr>
<td>65-74</td>
<td>2 (15)</td>
</tr>
<tr>
<td>75-84</td>
<td>3 (23)</td>
</tr>
<tr>
<td>≥85</td>
<td>0 (0)</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Man</td>
<td>8 (62)</td>
</tr>
<tr>
<td>Woman</td>
<td>5 (38)</td>
</tr>
<tr>
<td><strong>Highest educational degree</strong></td>
<td></td>
</tr>
<tr>
<td>PhD</td>
<td>3 (23)</td>
</tr>
<tr>
<td>Master’s</td>
<td>4 (31)</td>
</tr>
<tr>
<td>Bachelor’s</td>
<td>3 (23)</td>
</tr>
<tr>
<td>College degree</td>
<td>1 (8)</td>
</tr>
<tr>
<td>High school</td>
<td>2 (15)</td>
</tr>
<tr>
<td>Less than high school</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

Table 1. Demographic information of the participants.
Evaluation Outcomes

In order to evaluate the GoodTimes app, we conducted a user study with 13 older adult participants to collect feedback on their experiences with the app and its usability. Table 2 encapsulates participants’ feedback on their conversational experiences with the VA. The outcomes revealed an overwhelmingly positive response to their interaction experience with the VA. A significant majority of participants (12 out of 13) endorsed the efficacy and relevance of the VA in their conversations. Specifically, 8 participants were in strong agreement that the VA disseminated pertinent information, while the remaining 5 concurred with this sentiment. Furthermore, a compelling majority (11 out of 13) expressed strong affirmation regarding the VA's accuracy in sharing memory-related information. In terms of the VA’s language clarity and appropriateness, 12 participants were either in strong agreement or in agreement, indicating near-universal approval.

The P value, a measure of statistical significance, provides robust statistical validation of these findings. For instance, the statement “The VA provided relevant information” had a P value of <.001, suggesting that the observed results are extremely unlikely to have occurred by chance alone. Similarly, the almost identical P values for “The VA provided correct information” and “The VA's language is appropriate and easy to understand” (both P<.001) underscore the authenticity and significance of these findings. A P value below the typical threshold of .05 indicates a significant difference from the expected neutral response. This provides strong evidence that participants genuinely felt the statements were accurate descriptors of their experiences.

Table 3 depicts a profound capacity to foster positive emotions and reminisce among the study participants. It was noteworthy that every participant either agreed or strongly agreed that the app invoked cherished memories. Moreover, a substantial majority signaled their agreement or strong agreement with the app’s efficacy in reviving memories of dear ones, such as friends and family. Impressively, a significant 11 out of 13 participants articulated that the app augmented their happiness.

In Table 3, P values offer compelling evidence of these perceptions. For the statement “Brings a lot of Good Memories,” a P value of <.001 indicates an exceptionally significant result, suggesting the overwhelmingly positive feedback was not a mere coincidence. The sentiment “accelerates thinking about friends and family” also received a P value of <.001, reinforcing the strong affirmation of the app’s ability to stir memories of loved ones. Furthermore, the feedback “Makes me happy” also manifested a P value of <.001, emphasizing that a significant number of participants derived joy from the app’s use. These P values, being well below the conventional .05 threshold, bolster the claim of the app’s potent capability to enhance emotional health through memory stimulation.

During the usability assessment phase of our survey, we gauged the app’s interface using standardized usability prompts. Table 4 depicts an overview of the results on app usability. The results elucidated a prevailing sentiment of approval among respondents concerning the app’s usability. Notably, a significant desire was expressed to engage with the system regularly, as indicated by a P value of .005, which suggests this sentiment was not merely by chance. Additionally, the system’s helpfulness and its design simplicity garnered significant endorsement, as evidenced by the compellingly low P values of <.01, respectively. This denotes a genuine appreciation for the system’s functionality and design among users.

### Table 2. Participants’ feedback on their conversational experience with the voice assistant (VA).

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree, n</th>
<th>Agree, n</th>
<th>Neutral, n</th>
<th>Disagree, n</th>
<th>Strongly disagree, n</th>
<th>Mean (SD)</th>
<th>2-tailed t (df)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conversation was pleasant</td>
<td>8</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4.3846 (0.9608)</td>
<td>5.1959 (12)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Conversation was fluent and natural</td>
<td>4</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4.0769 (0.8623)</td>
<td>4.5029 (12)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>The VA provided relevant information</td>
<td>8</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.6154 (0.5064)</td>
<td>11.5016 (12)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>The VA provided correct information</td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.8462 (0.3755)</td>
<td>17.7272 (12)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>The VA’s language is appropriate and easy to understand</td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.8462 (0.3755)</td>
<td>17.7272 (12)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

### Table 3. Participants’ feedback on their emotional response to app usage.

<table>
<thead>
<tr>
<th>Statement</th>
<th>Strongly agree, n</th>
<th>Agree, n</th>
<th>Neutral, n</th>
<th>Disagree, n</th>
<th>Strongly disagree, n</th>
<th>Mean (SD)</th>
<th>2-tailed t (df)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brings a lot of Good Memories</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4.7690 (0.4385)</td>
<td>14.5455 (12)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Accelerates thinking about friends and family</td>
<td>10</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4.6920 (0.6304)</td>
<td>9.6773 (12)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Makes me happy</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>4.1538 (0.8987)</td>
<td>4.6290 (12)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>
Another commendable finding was the respondents’ confidence in using the app without external technical assistance. This was statistically supported by a P value of .03, reflecting a valid level of user self-assurance. However, the topic of needing technical support to operate the system did receive varied responses. Interestingly, the statement “I think that I would need the support of a technical person to be able to use this system” had a P value of .49, indicating that this sentiment was not statistically significant and could likely be attributed to random variability.

**Discussion**

**Principal Results**

We designed, developed, and tested an interactive photo album app called GoodTimes that uses AI technology to engage in conversations with users and tell stories about pictures, including family, friends, and special moments. The app was developed using state-of-the-art AI technologies, including image recognition, natural language processing, knowledge graph, logic, and machine learning. We constructed a comprehensive knowledge graph that models the information required for effective communication, including photos, people, locations, time, and stories related to the photos. We then developed a VA that interacts with users by leveraging the knowledge graph and machine learning techniques.

In order to evaluate the GoodTimes app, we conducted a use-case study to verify its various functions in different real-life scenarios. Additionally, we conducted a user study with 13 older adult participants to collect feedback on their experiences with the app and its usability. We found that the feedback from our participants was highly positive, with 92% (12/13) reporting a positive experience conversing with GoodTimes. All participants mentioned that the app invoked pleasant memories and aided in recollecting loved ones, resulting in a sense of happiness for the majority (11/13, 85%). Additionally, a significant majority found GoodTimes to be helpful (11/13, 85%) and user-friendly (12/13, 92%). Most participants (9/13, 69%) expressed a desire to use the app frequently, although some (4/13, 31%) indicated a need for technical support to navigate the system effectively.

**Limitations**

Our research has some limitations that we aim to address in future work. First, we plan to increase family members’ active involvement in the process of using the app by making it easy and enjoyable for them to upload pictures and record their voices. Additionally, the findings may not be generalizable to both cognitively impaired and cognitively intact older adults due to the small convenience sample used in this study.

To better understand the potential of our app as a supplementary tool for reminiscence therapy, we plan to deploy the app to their homes or assisted living environments, allowing them to use it for an extended period of time. By doing so, we can observe the app’s impact over time and gain valuable insights into how it can provide emotional and mental stimulation to improve their quality of life.

**Comparison With Previous Work**

Previous research has shown that reminiscence therapy, which involves the use of photos and videos to stimulate long-term memory, can improve the emotional well-being of older adults [5,37]. However, providing personalized reminiscence therapy can be challenging for caregivers and family members. Various studies have investigated technology as a means of supporting reminiscence therapy for older adults. For instance, Chen et al [38] developed an app that used a lifelogging device to capture photos and videos and presented them in a timeline format, resulting in improved quality of life and cognitive function for participants. However, this approach is limited to recent events and can be inconvenient for users to wear many devices. Additionally, the mashup process requires significant caregiver involvement.

Another study by Tsao et al [39] developed an augmented reality app that allows users to interact with virtual versions of their memories. The app was found to increase participants’ sense of control and satisfaction with their lives. The study by Schoneveld [40] developed an AR photo album prototype to facilitate communication between a person with dementia and their caregiver, family member, or friend, aiming to evoke more details and elements of memory and contribute to additional discussion material. The prototype has shown positive results in low- and high-fidelity prototype testing with experts and proxy testers. Another study [41] created digital reminiscence
and music therapies using prompts such as photos, videos, and music. The study focused on the rural population, which has reduced access to dementia care services.

Compared with these studies, our GoodTimes app uses AI technology to engage in conversations with users and tell stories about pictures, including friends, family members, and special moments. This personalized approach aims to provide a more engaging and emotionally supportive experience for older adults, requiring minimum efforts from caregivers. This study’s results showed that the app was well-received by participants, who found it helpful, easy to use, and enjoyable. In conclusion, while previous research has shown the benefits of reminiscence therapy for older adults, our GoodTimes app adds a new dimension to the field by using AI technology to provide a personalized and engaging reminiscence therapy experience.

Conclusions
In conclusion, this study demonstrated the potential of the GoodTimes app to provide personalized reminiscence therapy to older adults, improving their emotional well-being. The use-case study and user study results showed that the app was well-received by participants and provided a helpful, easy-to-use, and enjoyable experience. Although this study has some limitations, such as the need for a larger sample size and a longer evaluation period, we plan to address these limitations in future work. Our GoodTimes app adds a new dimension to the field of reminiscence therapy by using AI technology to provide a personalized and engaging experience. Overall, we believe that the GoodTimes app has the potential to positively impact the lives of older adults and their families.

Acknowledgments
The authors would like to thank all the interviewees for their contributions to the study. This work was supported by the National Science Foundation (NSF) with award number 2218046.

Conflicts of Interest
None declared.

References


Abbreviations

AI: artificial intelligence
DM: dialogue management
GPT: Generative Pre-trained Transformer
IIM: intent identification model
SUS: System Usability Scale
VA: voice assistant